La crise des fondements : quelle crise ?

Actes de l'atelier tenu du 7 au 9 octobre 2011
Sous la direction de François Lepage et Karine Fradet

Les Cahiers d'Ithaque
www.revueithaque.org/cahiers

Actes de l'atelier

La crise des fondements : quelle crise ?

Les Cahiers d'Ithaque

La crise des fondements : quelle crise ?

Sous la direction de
François Lepage et Karine Fradet

Les Cahiers d'Ithaque

La crise des fondements : quelle crise ?

Actes de l'atelier
7 au 9 octobre 2011
Université de Montréal

Sous la direction de
François Lepage et Karine Fradet

L'équipe de la revue Ithaque tient à remercier cordialement les participants à l'atelier organisé par François Lepage et Karine Fradet, tenu à l'Université de Montréal du 7 au 9 octobre 2011. Nous tenons également à remercier le Conseil de recherches en sciences humaines (CRSH) pour avoir rendu possible l'édition des actes de cet atelier.

Revue Ithaque
Département de philosophie
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal, Québec, H3C 3J7

www.revueithaque.org/cahiers
revue.ithaque@gmail.com

ISBN 978-2-9814181-0-4

Dépôt légal - Bibliothèque et Archives nationales du Québec (2013)
Dépôt légal - Bibliothèque et Archives Canada (2013)

Imprimé par Le Caïus du livre inc. (Montréal)

Table des matières

Avant-propos

Karine Fradet et François Lepage

L'expression même de « Crise des fondements » est relativement ambiguë. Elle renvoie d'une manière globale à une période qui débute vers 1870, époque où certains mathématiciens se posaient des questions philosophiques sur les géométries non euclidiennes, l'existence des nombres complexes, la nature des nombres réels et, un peu plus tard, le rôle fondateur de la théorie des ensembles. S'il y a un consensus relatif sur la période où débute cette crise, la date de sa fin, elle, est assez imprécise et dépend des multiples positions philosophiques.

La nature même de la crise varie selon les auteurs. Pour certains, la crise concerne essentiellement l'apparition de paradoxes comme ceux de Burali-Forti, Cantor ou Russell et elle prend fin avec l'apparition de la théorie des types et surtout avec l'apparition d'axiomatiques comme celle de Zermelo-Fraenkel, deux approches bien différentes qui évitaient les « cercles vicieux », selon l'appellation de Russell. Ces cercles vicieux permettaient de créer des monstres contradictoires comme l'ensemble de tous les ensembles, qui possède une cardinalité plus grande que lui-même, ou encore l'ensemble de tous les ensembles qui ne s'appartiennent pas, qui possède l'étrange propriété de s'appartenir si et seulement si il ne s'appartient pas.

Cette conception de la crise, dont la résolution se fera autour de 1908 avec la publication de « Set theory as based on the theory of types » par Russell et presque simultanément de « Untersuchungen über die Grundlagen der Mengenlehre I » par Zermelo, n'émeut pas particulièrement le monde des mathématiciens. Poincaré, qui n'était pas le moindre d'entre eux, avait déclaré à propos de la contradiction de Russell : « La logique n'est plus stérile, elle engendre des contradictions. » Nombreux étaient ceux qui pensaient que les contradictions provenaient d'un mauvais usage du langage, de l'utilisation de définitions imprédicatives, de définitions d'objets qui utilisent des totalités qui contiennent déjà l'objet à définir. Cette crise a fait couler beaucoup d'encre, mais elle fait maintenant définitivement partie de l'histoire de la philosophie de la logique et des mathématiques.

Une autre controverse, considérée par plusieurs comme la véritable crise des fondements, concerne la nature même des mathématiques et renvoie à la querelle philosophique entre Hilbert et Brouwer, entre les formalistes et les intuitionnistes. Qu'est-ce que l'existence d'un objet mathématique? Ou encore plus fondamentalement, qu'est-ce qu'une vérité mathématique? Ce clivage philosophique a encore, nous pourrons le constater à la lecture de ce recueil, beaucoup d'échos aujourd'hui.

Enfin, il y a ce qu'on pourrait appeler le « laissé pour compte » de la crise, le système de Leśniewski, apparu trop tard sans doute, mais qui continue de hanter, par sa beauté et sa rigueur, de nombreux philosophes qui persistent à s'intéresser à la crise des fondements.

Voici une brève présentation des textes de ce recueil. Dans « Ernst Schroeder and Zermelo's Anticipation of Russell's Paradox », Bernard Linsky revient sur un épisode bien connu, mais peu discuté dans la littérature : dans une lettre à Husserl, Zermelo fait remarquer qu'il n'y a pas d'ensemble de tous les ensembles qui s'appartiennent à eux-mêmes. Ce paradoxe prendrait son origine dans la critique que Husserl lui-même a faite d'un argument de Schroeder dans son *Algebra der Logic* de 1890. C'est une anticipation de ce que l'histoire retiendra comme étant le fameux paradoxe de Russell (l'ensemble de Zermelo est en fait le complément de l'ensemble contradictoire de Russell). Suivant un argument dû à Wiener, Schroeder aurait anticipé la théorie des types de Russell. Il n'y aurait pas de trace de la réaction de Russell à qui Wiener aurait pourtant fait part de son interprétation.

Michèle Friend, dans « Embracing the crisis in the foundation of mathematics », soutient que la principale leçon à tirer de la crise des fondements est qu'il n'y a pas un fondement des mathématiques, mais plusieurs : nous devons aller vers un pluralisme des fondements des mathématiques. Le point de départ est que le Pluralisme, dans la perspective des fondements des mathématiques, a pour thème central que la notion de vérité n'a de sens qu'à l'intérieur même d'une théorie mathématique. Michèle Friend soutient une certaine forme d'agnosticisme vis-à-vis de la question de savoir s'il y a un et un seul « fondement ». Il y a donc plusieurs théories fondationnelles qu'elle appelle « théories parapluie plausibles ». L'une d'entre elles est ZFC+ (quoi que ce « + » veuille bien signifier) qui est une sorte de lingua franca que l'on peut accepter ou rejeter. Elle conclut que le pluralisme n'est pas une position intermédiaire, mais une position tolérante et ouverte.

Le texte de Mathieu Bélanger, « La méthode axiomatique dans la crise des fondements », porte sur l'évolution de la notion d'axiome de Frege à Noether. Mathieu Bélanger nous rappelle d'abord la position de Frege pour qui les définitions ne sont pas créatives : elles ne font que nommer des concepts qui préexistent aux définitions. Il nous rappelle ensuite quel statut les axiomes de l'arithmétique ont pour Frege, soit celui de vérités évidentes qui n'appellent à aucune justification. À l'opposé, pour Hilbert, les axiomes ont un rôle, celui de formaliser nos intuitions. Les notions fondamentales sont celles de cohérence et d'indépendance des axiomes. Les axiomes ne sont pas fondationnels. Mathieu Bélanger s'attaque ensuite aux divergences profondes qui séparent la conception frégéenne de la conception hilbertienne et nous montre que les deux points de vue sont inconciliables. Enfin, l'auteur nous rappelle comment les travaux d'Emmy Noether dans les années vingt introduisent une nouvelle façon d'utiliser l'approche axiomatique pour démontrer des résultats totalement originaux sur la factorisation unique.

Héritier du logicisme de Frege et Russel, Carnap se lança dans le projet d'établir les fondements logiques des probabilités, ce que Karine Fradet présente dans « La fondation logique des probabilités chez Carnap ». Pour Carnap, la logique des probabilités est la logique inductive. Selon cette interprétation, une probabilité décrit une relation logique, soit le degré de confirmation d'une hypothèse selon des données observées. Il enrichit ensuite son système et démontre qu'il existe une infinité de méthodes inductives selon le poids que l'on accorde au facteur logique et au facteur empirique.

Si la crise des fondements a secoué la logique et la théorie des ensembles, elle n'a pas affecté le finitisme arithmétique de Kronecker qui généralise la décomposition euclidienne en nombres premiers aux polynômes irréductibles à coefficients algébriques. C'est la thèse que défend Yvon Gauthier dans « De Kronecker à Gödel via Hilbert. Les fondements arithmétiques et une crise sans fondement ». Yvon Gauthier montre la continuité qu'il y a entre le finitisme arithmétique de Kronecker et le finitisme de l'arithmétique intuitionniste de Gödel dans son interprétation fonctionnelle en passant par Hilbert qui voulait assurer le passage de l'arithmétique aux éléments idéaux de l'analyse. Le fondement des mathématiques dans l'arithmétique générale « ne pourra que nous rasséréner dans l'analyse critique de fondements qui ne connaissent pas de crise. »

Benoît Potvin, pour sa part, se penche sur un texte méconnu de Turing, *Systems of logics based on ordinals*, écrit en 1939 (c'est en fait sa thèse de doctorat dirigée par Church). L'idée de Turing est en quelque sorte de sauver le programme de Hilbert, profondément mis à mal par le théorème d'incomplétude de Gödel, en élaborant une logique ordinale. Il s'agit d'une suite de systèmes indexés sur les ordinaux, chacun d'eux étant incomplet, mais plus complet que les précédents. Benoît Potvin nous montre en quoi ce texte est fondamental pour comprendre la complexité d'une fonction calculable, concept fondamental en informatique théorique.

La contribution de Clayton Peterson, « La portée épistémique de l'axiomatisation de la physique chez Hilbert », sort du cadre strict de la problématique de la crise des fondements. Elle porte sur l'influence que la méthode axiomatique développée par Hilbert pour fonder les mathématiques a eue à l'extérieur du champ des fondements des mathématiques, en particulier en physique. Après avoir remarqué que la méthode axiomatique de Hilbert permet de montrer que certaines conclusions ne peuvent formellement être déduites des hypothèses, il souligne que cette méthode permet de recentrer le travail sur la recherche d'autres hypothèses. Clayton Peterson défend ainsi une interprétation épistémique du formalisme à la Hilbert.

Dans « Preuves intuitionnistes touchant la première philosophie » Joseph Vidal-Rosset montre que l'intuitionnisme possède des racines profondes dans l'histoire de la philosophie. Il entreprend en effet de démontrer, à la suite de Vuillemin, que Descartes était un intuitionniste avant la lettre et que la logique sous-jacente aux *Méditations* est la logique intuitionniste et même que les deux preuves fondamentales des *Méditations métaphysiques* peuvent être formalisées en déduction naturelle du calcul des prédicats intuitionniste (c'est-à-dire sans élimination de la négation).

Le texte de Pierre Joray, « A non reductionist logicism with explicit definitions », s'inscrit dans la tradition issue des travaux de Leśniewski. Utilisant la notion de définition explicite telle qu'on la retrouve dans l'*Ontologie* de Leśniewski et introduisant un axiome de l'infini, Pierre Joray montre que l'on peut reconstruire l'arithmétique dans le cadre de l'*Ontologie*. Toutefois, cette réduction a non pas une portée ontologique (au sens habituel de ce mot), mais une portée épistémique : l'usage exclusif de définitions explicites montre l'existence d'une conception purement analytique de la connaissance de l'arithmétique.

Enfin, François Lepage montre dans « Ontologie et théorie des ensembles » que, contrairement à ce que la plupart croient, la théorie des ensembles et l'*Ontologie* de Leśniewski ne sont pas incompatibles. Bien au contraire, il reconstruit, en utilisant des définitions adéquates, les axiomes de Zermelo-Fraenkel à l'intérieur de l'*Ontologie* et démontre qu'ils sont valides.

Ernst Schroeder and Zermelo's Anticipation of Russell's Paradox

Bernard Linsky
University of Alberta

Abstract

Ernst Zermelo presented an argument showing that there is no set of all sets that are members of themselves in a letter to Edmund Husserl on April 16th of 1902, and so just barely anticipated the same contradiction in Betrand Russell's letter to Frege from June 16th of that year. This paper traces the origins of Zermelo's paradox in Husserl's criticisms of a peculiar argument in Ernst Schroeder's 1890 Algebra der Logik. Frege had also criticized that argument in his 1985 "A Critical Elucidation of Some Points in E. Schroeder Vorlesungen über die Algebra der Logik", but did not see the paradox that Zermelo found. Alonzo Church, in "Schroeder's Anticipation of the Simple Theory of Types" from 1939, cricized Frege's treatment of Schroeder's views, but did not identify the connection with Russell's paradox.

Bertrand Russell wrote to Gottlob Frege on June 16th, 1902, saying that he had found "one point where I have encountered a difficulty" in reading over *Grundgesetze der Arithmetik* in preparation for publishing his own *Principles of Mathematics*. The difficulty is described as follows:

> Let *w* be the predicate: to be a predicate that cannot be predicated of itself. Can *w* be predicated of itself? From each answer its opposite follows. Therefore we must conclude that *w* is not a predicate.

Russell immediately follows this with what sounds like the different, and more familiar, problem of the class of all classes which do not belong to themselves, but is presented as another statement of the same point of difficulty:

> Likewise there is no class (as a totality) of those classes which, each taken as a totality, do not belong to themselves. From this I conclude that under certain circumstances a definable collection [Menge] does not form a totality.[1]

These are in fact two different problems. The first involves the notion of a predicate that is or is not predicated of itself. As Frege pointed out immediately in his response to Russell, if one makes a sharp distinction between predicates that are true of individuals, and the (higher order) predicates that are true of those predicates, as Frege himself did within his theory of concepts and objects, then it simply won't make sense to think of a predicate applying to itself. There will be no predicate "to be a predicate that is not predicated of itself," and so the paradox cannot arise. In fact, at the time, Russell did not accept the notion that concepts would be distinguished by type, and only acknowledged that something of the sort might be necessary in Appendix B of *Principles of Mathematics*. This first paradox was the most significant for Russell. He only formulated the second when he saw that related versions could be constructed for other theories, such as Frege's. Frege immediately acknowledged that there was a problem for his system in the second paradox.

Frege responded to Russell's letter quite quickly, with a reply dated June 22nd, just six days later. Frege concludes as follows:

> It seems, then, that transforming the generalization of an equality into an equality of courses-of-values (§9 of my *Grundgesetze*) is not always permitted, that my Rule V (§20, p.36) is false, and that my explanations in §31 are not sufficient to ensure that my combinations of signs have a meaning in all cases.

[1] Van Heijenoort (1967, pp. 127-128).

Russell did come to the paradox by considering Cantor's theorem that the set of subsets of a given set, its power set, cannot be put into a one-to-one correspondence with that set by considering the hypothesis that there is a universal set. If one follows the proof of Cantor's theorem the paradoxical set of all sets that are not members of themselves is the set that is added in the power set of the universe. Russell himself, however, did not see this as a problem for Cantor's set theory as much as for his own. Russell originally found the paradox for his own views in *Principles of Mathematics*, in which classes are represented by "class concepts", which in fact denote the members of a class. The members of the class can be taken "as one" or "as many" and it is when we speak of a class as one that it makes perfect sense for it to be an instance of a class, and so the concept "is not predicated of itself" can perfectly well be applied to itself (as one), thus leading to the contradiction.

In this paper, however, I will follow the earlier history of the second version of the paradox, which was anticipated by Zermelo. The paradox developed out of an argument of Ernst Schröder, which was subsequently discussed by Frege, Edmund Husserl, and it finally appeared in a letter from Zermelo to Husserl, although Hilbert also had his own version. Finding that this version of the paradox did in fact have a long history that preceded Russell's letter to Frege, if anything, will only add to the argument that it was that first version that most immediately interested Russell, the problem for his own theory, and that he wrote to Frege only to show how a similar problem arose in Frege's *Grundgesetze*.

Any discussion of anticipations of the second version of the paradox might look like an intervention in a mathematicians' *Prioritätstreit* (priority dispute), but it in fact leads to an interesting series of events in the history of logic, pieces of which are familiar, but which, it seems, has not been formed into a single story before.

The most prominent claim of priority is in Ernst Zermelo's famous 1908 paper "A new proof of the possibility of a well-ordering." Zermelo adds a footnote (numbered 9) to a mention of Russell's discussion of his "set-theoretic antinomy" in *Principles*:

> … I had, however, discovered this antinomy myself, independently of Russell, and had communicated it prior to 1903 to Professor Hilbert among others.

In a letter of 1903, thanking Frege for a copy of the second volume of his *Grundgesetze* with its discussion of Russell's paradox, Hilbert told Frege that he had heard of the paradox from Zermelo some years before, confirming Zermelo's claim. In fact, Hilbert had his own version of the contradiction, concerning the class of all sequences of "self-mappings'" of the numbers onto the numbers, using a diagonal argument to show that the set of such self-mappings does not exist.[2] In his biography of Zermelo, Ebbinghaus claims to just be following the "custom" at Göttingen, where the paradox was attributed to Zermelo alone. (Ebbinghaus actually more closely follows Abraham Fraenkel's terminology and calls it the "Zermelo Russell paradox" in the biography.[3])

"Hilbert's Paradox" is similar to Zermelo's in proving that some set does not exist, but in his case it is a set that one would expect given certain standard set theoretic operations. The paradox can be reconstructed from notes on Hilbert's lecture from 10 July 1905.[4] Hilbert derives a contradiction by taking the set of mappings of the numbers into the numbers M and then using two principles of set theory. The first principle allows one to unite "several sets and even infinitely many into a union." The other asserts that "in every case well-defined sets arise from well-defined sets by the self-mapping operation." Thus from the set M we can formulate the set M^M of all mappings from M into M. Hilbert then considers the set U derived by applying these "operations of addition and mapping an arbitrary number of times." Finally, he applies the mapping principle one more time, to U, so $F = U^U$. Now it would appear that F is already a subset of U, but, by a familiar diagonal argument based on Cantor's theorem, we can show that there is an element in F not in U.

---

[2] Ebbinghaus (2007, pp. 45-47) and Peckhaus (2004, pp. 505-506).

[3] Ebbinghaus (2007, pp.46-47) and Fraenkel (1927). Fraenkel, perhaps the first to use this term in print, puts "Zermelo" in parentheses, and calls it the "(Zermelo) Russell" paradox once in his presentation and not again in his discussion. He does not give any reason for his usage.

[4] Ebbinghaus (2007, pp. 45-47).

Hilbert's words are not precise; in particular it isn't clear what he means by applying the two operations an "arbitrary number of times." One guess is that he thinks of F as being the union of an infinite number of sets:

$$M^M,\ (M^M)^M,\ M^M \cup (M^M)^M,\ ((M^M)^M)^M),\ M^M \cup (M^M)^M \cup ((M^M)^M)^M),\ \ldots$$

It appears that, if we have continued this process an "arbitrary number of times", then anything in U^U must have been already included in this process and so the process yields nothing more than U. Of course we immediately sense that the "arbitrary" number of operations must pass through all the ordinals, taking a union after ω many iterations of M and beginning again with another process of either union or exponentiation for each ordinal number, and so on. Seen this way "Hilbert's Paradox" is a proof that there is no set of all ordinals. Hilbert was, reportedly, suspicious of the "philosophical" notions used in other versions of the paradox, such as the notion of a "set of all sets" or even a "set of all ordinal numbers", and so preferred to limit himself to more familiar mathematical notions, such as mapping, or the application of an operation an "arbitrary" number of times. Thus, the argument can be seen, in Hilbert's way, as a contradiction which can be derived from two apparently ordinary mathematical operations, considering function spaces and taking unions of arbitrary collections of things already constructed. It doesn't point to an internal contradiction in the very notion of set or class. It may well be an "antinomy", as Zermelo puts it, a real internal contradiction, but it is a contradiction in a theory, one which Hilbert held should be prevented by constructing a (provably) consistent axiomatic theory of sets.

Zermelo sent a letter on April 16th, 1902, to his former teacher, Edmund Husserl. Husserl's notes on the letter were found among the papers in the Husserl archives thirty years ago.[5] There, Zermelo reports a result that he had obtained some years before.[6] The letter was occasioned by a review of Schröder's *Algebra der Logik* (1890) that

5 Rang and Thomas (1981) and Ebbinghaus (2007, p. 46).

6 See Husserl (1979 p. 399) for Husserl's notes on this letter.

Husserl had written in 1891.[7] Schröder had presented a proof that a universal class, one that contains "everything conceivable", leads to a contradiction. (It thus appears that Schröder joined Cantor in being among the first to claim that there is a concept which does not have an extension.)

One primary goal of Schröder's lectures was to promote the work of Charles S. Peirce "and his school" among German logicians (Schröder 1890, p.iii). In Lecture IV, devoted to the theory of classes, Schröder presents the algebraic account of classes using one primitive notion of "subsumption" (*Subsumtion*). The assertion that a class *a* subsumes a class *b*, symbolized as *a* € *b* which is read as "*a* is *b*" or "all *a* are *b*", clearly means what we would express as "*a* is a subset of *b*", symbolized as $a \subseteq b$.[8] This logic interprets all predications to be of this form "*a* is *b*", relating one class to another.

One part of Schröder's account of sets is a rejection of Boole's notion of the *universe of discourse*, which in this framework will be the element 1 in the algebra of classes. The argument below is from page 245 of Schröder (1890) as quoted by Frege:

> As we have laid down, 0 would have to be contained in *every class* that can be got out of the manifold 1; … 0 would have to be the subject of *every* predicate. Now suppose we took *a* to be *the class of those classes of the manifold that are equal to* 1 (which would certainly be permissible if we could bring everything thinkable into the manifold 1), then this class of its very nature contains just one class, viz. the symbol 1 itself, or alternatively the whole of the manifold, which constitutes the reference of the symbol; *but therefore besides this it would contain "nothing,"* i.e. 0. Hence 1 and 0 would make up the class of the objects that are to be equal to 1; and so we should have to admit not only: 1 = 1 but *also*: 0 = 1. For a predicate that applies to a class -- in our case, the predicate: to be identically equal to 1 -- must also

[7] See Husserl (1979 p. 36).

[8] Actually the symbol is the identity sign = with a left parenthesis (overwritten, but it is close enough to the readily available Euro sign to justify that as the symbol for any revival of interest in Schröder's logic.

> apply to every individual in the class, by Principle II. (Frege, 1895, p.91)

For Schröder, all predications are assertions of subsumptions, and in particular, the predication 'is equal to 1', which we would express as '$x = 1$', becomes an assertion about subsumption. Given that the empty set, represented by 0, is subsumed by every class *a*, i.e. 0 € *a*, it is also subsumed by that of the things equal to 0, hence we derive 0 = 1, a contradiction.[9] This he takes to be a proof that there is no absolutely universal class 1. There will be sets that are not subsumed under 1, in particular, the empty class 0.

Husserl charged in his review that Schröder had ignored the distinction between subsets (the notion of a "subordinate class") and members. While it is true that the empty class 0 is a subset of every set, it is not a member of every set. In particular that 0 is a subset of the set of entities equal to 1 does not imply that it is equal to 1, the supposed contradiction which is derived from the assumption that 1 is a set of all sets. Zermelo then wrote to Husserl to point out that "on the issue, not the method of proof, Schröder is right…"[10]. Translated from the original German *Gabelsburger Stenographie* shorthand, the notes describe the argument as follows:

> A set *M*, which contains each of its subsets *m*, *m'*, … as elements, is an inconsistent set, i.e., such a set, if at all treated as a set, leads to contradictions.
>
> PROOF. We consider those subsets *m* which do not contain themselves as elements.

---

[9] This is a charitable reconstruction. Notice that Schröder argues that the class of things equal to 1 includes 1 "and nothing else", i.e., also the empty class ("nothing"). This sounds like Carnap's later interpretation of what Heidegger says about "*Das Nichts*", or "Nothing", as something which is included in classes. Frege criticizes Schröder's argument on this point (p.98), anticipating Carnap's criticism of Heidegger: "If we say that the class *a* contains nothing besides the Moon, then we are denying the proposition that the class contains something besides the Moon; but we are not thereby asserting that the class contains, besides the Moon, an object with the name 'nothing'."

[10] Rang and Thomas (1981, p. 16).

> (M contains as elements each of its subsets; hence subsets of M will also contain certain subsets as elements, themselves [not] being elements, and now we consider just those subsets m, which may perhaps contain other subsets, but not themselves as elements.)
> These constitute in their totality a set M_0 (i.e., the set of all subsets of M which do not contain themselves as elements), and now I prove of M_0,
>
> (1) that it does not contain itself as an element,
> (2) that it contains itself as an element.
>
> Concerning (1): M_0, being a subset of M, is itself an element of M, but not an element of M_0. For otherwise, M_0 would contain as an element a subset of M (namely, M_0 itself) which contains itself as an element, and that would contradict the notion of M_0.
>
> Concerning (2): Hence M_0 itself is a subset of M which does not contain itself as an element. Thus it must be an element of M_0.[11]

So stated this is a proof that no set contains all of its own subsets as members. A universal set of all things, however, would certainly include all of its subsets as members, as those are all sets of things. The set M_0, then, the set of all subsets of the universal set that do not contain themselves, will thus be simply the set of all sets that do not contain themselves. M_0, in this case, is the Russell set. The proof that M_0 leads to contradiction is the same argument that Russell gives, if M_0 is a member of itself then it is not, and if it is not a member of itself then it is. We have the same contradiction as in Russell's letter, and both, it seems, come fairly directly from applying Cantor's theorem to a set of all sets (or in Zermelo's case, a set which contains at least all its subsets.)

But do we have Russell's paradox? In fact, what we have is a theorem to the effect that there is no set which contains all of its subsets as members. This is a theorem about sets. Indeed in another

[11] Rang and Thomas (1981, pp. 16-17).

paper from 1908, "Investigations in the foundations of set theory", we find it openly listed as a theorem with the following proof:

> 10. Theorem. Every set M possesses at least one subset M_0 that is not an element of M. Proof. It is definite for every element x of M whether $x \in x$ or not; the possibility that $x \in x$ is not in itself excluded by our axioms. If now M_0 is the subset of M that, in accordance with our Axiom III [Zermelo's Axiom of Separation], contains all those elements of M for which it is not the case that $x \in x$, then M_0 cannot be an element of M. For either $M_0 \in M_0$ or not. In the first case, M_0 would contain an element $x = M_0$ for which $x \in x$, and this would contradict the definition of M_0. Thus M_0 is surely not an element of M_0, and in consequence M_0, if it were an element of M, would also have to be an element of M_0, which was just excluded.

Zermelo concludes the proof with this remark:

> It follows from the theorem that not all objects x of the domain B can be elements of one and the same set; that is, *the domain B it is not itself a set*, and this disposes of the Russell antinomy so far as we are concerned.[12]

So Zermelo "disposes" of Russell's antinomy by presenting it as a theorem that a certain collection is not a set, a proof by a *reductio ad absurdum* argument.[13] Burali-Forti's proof from 1897 that the ordinals cannot be well-ordered was also taken as such an argument.[14] If every

[12] Van Heijenoort (1967, p. 203).

[13] In particular, there is no set that contains all of its subsets as members. This might sound a bit like the Axiom of Foundation, which says that there is no non-empty set that overlaps with every one of its members, and so shares a member with every one of its subsets. They are clearly very different claims, but do sound alike in the sense of saying that there is no set of a given sort.

[14] See Moore and Garciadiego (1981) for an account of how Burali-Forti's original theorem came to be seen as a "paradox". It started life as a claim that the set of all ordinals could not be well-ordered. After Burali-Forti's

set can be well ordered then it follows that there is no set of ordinals. Indeed Zermelo's proof would seem to be also a proof related to the idea of "absolute infinities," or of classes that are in some sense "too big" to be sets. A "set of all sets" would certainly contain all of its subsets as members, and so one immediate result from Zermelo's theorem is that there is no set of all sets.

Zermelo considers that he has "disposed" of Russell's paradox by showing it to be a simple non-existence proof, although a surprising one. But there is no "antinomy" in that. A proof that there is no set with a certain description is different from proving that nothing satisfies a given description, so that the set of things satisfying the description is empty. It is really a proof that some expression, which seems to refer to a set, in fact doesn't, because there is no such set. For every counter-example to the unrestricted comprehension principle, that is, a case of a predicate without a set as its extension, there will surely be a corresponding proof that there is no such set, that is, that no set that contains just the objects which satisfy that predicate. But there is no trivial reformulation of every non-existence theorem into a counter-example to the unrestricted comprehension principle. To think that would be to make a logical mistake. Suppose one thought that "there is no set *y* which contains all of its (*y*'s) own subsets" directly provides a predicate which does not correspond with a set, thus a counter-instance to the unrestricted comprehension principle:

$$\exists y \forall x \, (x \in y \equiv \ldots x \ldots)$$

Such an instance would have to say that there exists a *y* such that any *x* is an element of *y* if and only if … what? If and only if *x* is a subset of *y*? That would violate the clause in the comprehension principle stated above which bans free occurrences of *y* in the formula … *x* … that determines membership in the putative set. It is not possible to formulate the claim that no set contains all its own subsets as members as a claim that there is no set of all things

incorrect definition of the relevant notion of well-ordering was corrected, the result was not seen as a paradox until Russell first so presented it in Principles of Mathematics §301.

satisfying a certain formula.[15] So Zermelo does indeed propose that Russell should be seen as only proving a non-existence theorem rather than as proposing a counter-example to a comprehension principle that was presumably implausible from the beginning. There might be some surprise in discovering that there is no set of all sets and, in fact, that no set can contain all of its subsets as members, but it is not clear that that should count as a "paradox." In fact, Zermelo himself preferred the term "antinomy" for the result rather than paradox.[16] He says that "paradox" means "… a statement contradicting *the common opinion*; it doesn't contain anything of the *inner contradiction* as in the case for the paradoxes of Russell and Burali-Forti, and expressed by the term 'antinomy'." On the other hand, that an "antinomy", as Zermelo uses the term, can be derived within a formal theory is simply a proof that one of its axioms is incorrect and must be discarded. It doesn't mean necessarily that the very concepts or basic notions of a theory are at fault and lead to the contradiction. So far, then, it appears that while Zermelo had indeed anticipated the very mathematical argument that Russell later included in his letter, he did not see that it was a "paradox" which might affect several approaches to sets, including Frege's.

There is more to be said about Schröder's argument, however.[17] In fact, Frege discusses the argument in his essay "A critical elucidation of some points in E. Schröder's *Vorlesungen über die Algebra der Logik*" (1895). Frege begins his discussion with the quote above. He then continues as follows:

> On p. 246 the author shows that we can apply these considerations to any class *b* of the manifold, instead of 1, and thus reaches the conclusion $0 = b$.

---

[15] It is possible to present the proof that there is no set of all things as a counter-example to an instance of the comprehension principle: $\exists y \forall x\, (x \in y \equiv x = x)$, but this can't be done for every non-existence theorem.

[16] See Peckhaus and Kahle (2002, p.158). He says this in correspondence with Leonard Nelson in response to a paper about the "paradoxes" of Russell and Burali-Forti. Peckhaus and Kahle cite a postcard to Nelson, postmarked 22 December 1907, in the Nelson Papers in Bonn.

[17] I am grateful to Allen Hazen who reminded me that the Church and Frege papers explicitly discuss Schröder's argument.

> This contradiction comes like a thunderbolt from a clear sky. How could we be prepared for anything like this in exact logic! Who can go surety for it that we shall not again suddenly encounter a contradiction as we go on? The possibility of such a thing points to a mistake in the original design. Herr Schröder derives from this the conclusion that the original manifold 1 must be so made up that, among the elements given as individuals within it, there are found no classes that, for their part, contain within themselves as individuals any elements of the same manifold. This expedient, as it were, belatedly gets the ship off the sandbank; but had she been properly steered, she could have kept off it altogether. It now becomes clear why at the very outset, in shrewd prevision of the imminent danger, a certain manifold was introduced as the theatre of operation, although there was no reason for this in the pure domain-calculus. The subsequent restriction of this field for our logical activities is by no means elegant. Whereas elsewhere logic may claim to have laws of unrestricted validity, we are here required to begin by delimiting a manifold with careful tests, and it is only then that we can move around inside it.

A few lines later:

> When Herr Schröder stipulates (p. 248), as regards the original manifold, that among the elements given as 'individuals' there shall be found no classes that, for their part, comprise within them as individuals any elements of the same manifold, he is obviously distinguishing the case where something is given as an individual belonging to a manifold or class, where something is comprised within a class as an individual, from the case where something is contained as a class within a manifold or class. Herr Husserl makes a similar distinction, in his review of Schröder's work, between the expressions 'a class contains something as an element' and 'a class contains something as a sub-class,' and by this he tries to remove the difficulty. (Frege 1895, p. 92)

Note that Zermelo's theorem 10, that "Every set *M* possesses at least one subset M_0 that is not an element of *M*", in effect *proves* that one must distinguish the notions of subset ($\subseteq$) and member ($\in$) for the subsets of a given set are distinct from its members. While "disposing" of Russell's paradox by showing that the argument proves a perfectly sound theorem about sets, Zermelo's proof also shows that Schröder's conception of sets is fundamentally flawed. While avoiding the paradox by relying on his own axiomatic formulation of set theory, Zermelo also finds an important lesson about the nature of sets in the argument that produces the paradox. Frege had also carefully studied Schröder's argument, but missed the consequence for his own theory. Zermelo's argument properly deserves to be called an "anticipation" of Russell's paradox and he discovered it by studying the very passage that Frege criticized in Schröder.

Schröder's contradictory conclusion that 0 = 1 is, of course, part of a *reductio ad absurdum* argument. Schröder does not draw the conclusion that every class must contain some classes that aren't members of that class, but rather that every class must contain only classes that in turn do not contain members of the original class. These will be the classes that serve as "elements" of the original class. Frege sees this as an *ad hoc* solution. It "gets the ship off the sandbank."[18] Frege sees laws of logic as having "unrestricted validity", and so presumably, the universal quantifier ranges over everything, without restriction.

In a paper delivered in 1939, but not published until 1976, Alonzo Church presents Schröder's proposal as an anticipation of the simple theory of types. While the "elements" of a given set *a* will not subsume any classes which are also members of *a*, they may well subsume members of some other class *b*. Church proposes that we might see *a* and *b* as belonging to different types, where *a* is one type higher than *b*, as the elements of *a* are all classes containing members from the type of *b*. Frege's opposition to these restrictions, and his insistence that the laws of logic are unrestricted is seen by Church as a repetition of his view that all objects, including the "courses of

[18] Frege (1895, pp. 339-340), discussed in Church (1976, p. 410).

values" which are Frege's sets, are of the same type, and fall within the range of the universal quantifier for objects.[19]

Frege cites the same review of Schröder that Zermelo corrects in his letter from 1902. There is no evidence that Zermelo had read Frege's paper. At the least it appears that both had read Husserl's review, and reacted to it. Frege and Husserl had both identified the conflation of membership and the subset relation in Schröder's notion of subsumption. Frege, however, also noticed, with Zermelo, that the point of the argument was to show that there is no universal set, or unrestricted "domain of discourse." Indeed, as Church notes, Frege was to fall for the very sort of paradox that he accuses Schröder of trying to avoid, namely one that follows from assuming that there is a class of all things. It is tempting to speculate that Zermelo was aware of this dispute among the logicians about the "universe of discourse" and saw Cantor's idea of "absolute" infinities and the non-existence of a set of all cardinal numbers, as a clearer account of these same issues. As a consequence his diagonal argument, mathematically the same as Russell's, would also have seemed to him to be an anticipation of the use to which Russell put his argument. But, to continue the speculation, of course, one would have to accuse Zermelo of not seeing the simple theory of types as an alternative resolution to the set theoretic "antinomy" he had discovered. But, in fact, Zermelo and Russell seem to have been working in different worlds in set theory; Zermelo within the tradition of Cantor's set theory at Göttingen, and Russell, as always, refining and abandoning his own earlier views.

There is no direct evidence that Russell studied Schröder's argument, or saw in it either the idea of a contradiction in the notion of a set of all sets, or an anticipation of the theory of types in its conclusion. Russell, did, however, study Frege's "A Critical Elucidation …" and made extensive notes on it as part of his preparation for adding "Appendix A: The Logical and Arithmetical Doctrines of Frege" to *The Principles of Mathematics* in the summer of 1902, the same preparation that led to his letter to Frege.[20] Nothing

[19] Church argues that Frege's theory of "Stufe" of concepts and concepts that apply to first level concepts, etc., is not a theory of types of *objects* in the sense in which a genuine theory of simple types of classes is.

[20] See Linsky (2004a and 2004b).

relevant to this argument appears in the notes or in the ultimate appendix.

However, there is evidence about Russell's general ideas about Schröder which comes from his interaction with Norbert Wiener in 1913.[21] In September 1913, Norbert Wiener, then just eighteen years old, visited Russell in Cambridge. He had just completed his PhD thesis at Harvard University, entitled *A comparison between the treatment of the algebra of relatives by Schröder and that by Whitehead and Russell.* Ivor Grattan-Guinness (1975) found several pages of comments by Russell and replies by Wiener, following a series of discussions the two had in September and October of that year. Wiener also attended Russell's lectures and some letters home report on the interchange.

Russell wrote about the discussions as well, including this in a letter to Lucy Donnelly from 19 October 1913:

> At the end of Sept. an infant prodigy named Wiener, Ph.D. (Harvard), aged 18, turned up with his father... The youth has been flattered, and thinks himself God Almighty – there is a perpetual contest between him and me as to which is to do the teaching. (Grattan-Guinness 1975, p.105)

Wiener brought a copy of his thesis, and he and Russell discussed it in a series of meetings, with an exchange of letters between meetings.[22] The thesis and letters include a number of points about Schröder's logic. The topic of the thesis was a defense of the merits of Schröder's logic in comparison with *Principia Mathematica*. Two passages singled out by Grattan-Guinness are particularly relevant to the issues introduced above:

---

[21] What follows is based on Grattan-Guinness (1975).

[22] This correspondence clarifies Wiener's often cited remark that Russell later was not particularly impressed with the reduction of relations to sets of ordered pairs that was Wiener's first published contribution to mathematics. (Wiener 1953, p. 191). In fact Russell had told Wiener that he did not identify relations with sets of "couples", as was common in the tradition of Schröder and Peirce that Wiener was defending. See Grattan-Guinness, 1975, p. 122.

> A major point of contrast between Schröder's and Russell's systems is that Schröderian individuals correspond to unit classes of Russellian individuals. Thus Schröder has no analogue to Russell's relation of membership of an individual to a class. But it does not matter, since Schröder has no need of such a relation, contrary to the opinion of Padoa that he conflated membership and inclusion, and to the opinion of Russell that all predecessors of Peano and Frege regarded membership as a special case of inclusion. (Grattan-Guinness 1975, p. 124)

Russell does not accept Wiener's claim that Schröder didn't need the distinction between membership and subset. In one of his replies, he asks "Have you any evidence that Schröder knew that there was a difference between Peter and the class whose only member is Peter?" to which Wiener replies by repeating that Schröder is simply "not concerned" with that distinction. (Grattan-Guinness 1975, p. 128)

Wiener credits Schröder with an anticipation of Russell's theory of types in his discussion of the argument about the universal class under discussion above:

> No individual in a manifold can itself be composed of a collection of other individuals of that manifold. Instead, classes of individuals belong to the first 'derived' (*abgeleitete*) manifold of the 'original' (*ursprüngliche*) one. Classes of classes of individuals belong to the second derived manifold and so on. This creates a hierarchy of types corresponding to Russell's theory. The difference is that, while Russell can use more than one type at once, Schröder can speak only of one type at a time. (Grattan-Guinness 1975, p. 128)

(This is the very same as Church's analysis of Schröder's position, only expressed twenty-six years earlier.) There is no record of Russell's reaction to this part of the thesis, but it is clear that Russell had read in Wiener an expression of the view that Schröder had anticipated the theory of types in response to an argument that there cannot be a universal class to which everything belongs. It appears that Russell did not think that he alone had seen that a diagonal argument like that in Cantor's proof would establish that there can be no universal set of all sets. Although he studied and cited Zermelo's papers from 1908 as well, there is also no record of any response to the claim that others had anticipated the paradox. Schröder's logic was so alien to Russell's that it is understandable that he did not find anticipations of his own theory of types in it. Russell does not cite any sources for the notion of types that appears first in Appendix A of *Principles of Mathematics*, although clearly it was Frege's notion of *Stufe* or the hierarchy of concepts, concepts of concepts, etc. that must have inspired it.

As with the anticipation of the theory of types, Russell was also close to Zermelo's version of the paradox in Schröder's arguments, but clearly didn't arrive at it through that route. Russell's own paradox was the first paradox of predicates. He saw that the paradox of the set of all sets that are members of themselves was a result of similar thinking. His interest in the first paradox made him blind to the possibility of the second paradox raised by Schröder's argument, which Zermelo had identified. But then Frege, who also read Schröder's argument carefully didn't see the argument either. Their thoughts were somewhere else.

References

CHURCH, A., 1976. Schröder's Anticipation of the Simple Theory of Types, *Erkenntnis*, vol. 10, pp. 407-411.

CANTOR, G., 1899. Letter to Dedekind. In van Heijenoort 1967, pp. 113-117.

EBBINGHAUS, H.-D., (with Volker PECKHAUS), 2007. *Ernst Zermelo: An Approach to his Life and Work*, Berlin: Springer-Verlag.

FRAENKEL, A. A., 1927. *Zehn Vorlesungen über die Grundlegung der Mengenlehre*, Teubner: Leipzig and Berlin.

FREGE, G., 1893. *Grundgesetze der Arithmetik*, Band I/II, Jena: Verlag Hermann Pohle.

FREGE, G., 1895. Kritische Beleuchtung einiger Punkte in E. Schröder's Vorlesungen über die Algebra der Logik, *Archiv für systematische Philosophie*, Vol.1 (1895), pp. 433-456. Translated as, A Critical Elucidation of Some Points in E. Schroeder, *Vorlesungen über die Algebra der Logik,* in McGuinness, pp. 210-28.

FREGE, G,. 1902. Letter to Russell, in van Heijenoort (1967), pp. 127-28.

GRATTAN-GUINNESS, I., 1975. Wiener on the Logics of Russell and Schröder. An Account of his Doctoral Thesis, and of his Discussion of it with Russell, *Annals of Science*, 32, pp. 103-132.

HUSSERL, E., 1979. *Aufsätze und Rezensionen (1890-1910)*, Bernhard Rang, ed., Husserliana XXII.

LINSKY, B., 2004a. Russell's Marginalia in his Copies of Frege's Works, *Russell: The Journal of Bertrand Russell Studies*, n.s. 24(1), pp. 5-36.

LINSKY, B. 2004b., Russell's Notes on Frege for Appendix A of *The Principles of Mathematics*, *Russell: The Journal of Bertrand Russell Studies*, n.s. 24(2), pp. 133-72.

MOORE, G. and GARCIADIEGO, A. 1981. Burali-Forti's Paradox: A Reappraisal of its Origins, *Historia Mathematica* 8, pp. 319-350.

PECKHAUS, V. 2004. Paradoxes in Göttingen, in G. Link, ed., *One Hundred Years of Russell's Paradox*, Berlin: de Gruyter, pp. 501-516.

PECKHAUS, V. and Kahle, R., 2002. Hilbert's Paradox, *Historia Mathematica* 29, pp. 157-175.

RANG, B. and THOMAS, W., 1981. Zermelo's Discovery of the "Russell Paradox", *Historia Mathematica* 8, pp. 15-22.

RUSSELL, B., 1902. Letter to Frege, in van Heijenoort (1967), pp. 124-25.

RUSSELL, B., 1903. *The Principles of Mathematics*, Cambridge: Cambridge University Press.

SCHRÖDER, E., 1890. *Vorlesungen über die Algebra der Logik (Exacte Logik)*, Vol. I, Leipzig: Tuebner. Reprinted Bristol: Thoemmes, 2001.

VAN HEIJENOORT, J., 1967. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Cambridge: Harvard University Press.

WEINER, N., 1953. *Ex-Prodigy: My Childhood and Youth*, MIT Press.

ZERMELO, E. 1908a. *A New Proof of the Possibility of a Well-Ordering*, in VAN HEIJENOORT (1967), pp. 183-98.

ZERMELO, E., 1908b. *Investigations in the Foundations of Set Theory I*, in VAN HEIJENOORT (1967), pp. 199-215.

Embracing the Crisis in the Foundations of Mathematics

Michèle Friend
George Washington University

Abstract

The crisis in the foundations of mathematics is a conceptual crisis. I suggest that we embrace the crisis and adopt a pluralist position towards foundations. There are many foundations in mathematics. However, 'many foundations' (for one building) is an oxymoron. Therefore, we shift vocabulary to say that mathematics, as one discipline, is composed of many different theories. This entails that there are no absolute mathematical truths, only truths within *a theory. There is no unified, consistent ontology, only ontology* within *a theory.*

Conceptual crises teach us that there is an instability in our preconceptions, theories or in our philosophical goals. The crisis in the foundations of mathematics is something we can regret, or something we can learn from. What I, and others, have learned is to abandon the goal of providing one formal mathematical theory as a foundation for the whole of mathematics.

In abandoning this goal, our first step is to embrace a moderate form of pluralism: 'Pluralism in Foundations', where we accept several foundations of mathematics. However, in seriously entertaining the idea of competing, or equally plausible, foundations in mathematics, we learn that the *very notion* of a formal theory being a 'foundation' for mathematics is unstable, and therefore that the very idea of 'foundation' (in mathematics) should be revisited.

This line of thinking takes us to the next step, where we open the door to a world of Pluralisms. From the doorway, we can now consider Pluralism in: epistemology, ontology, methodology and truth. In this paper, I shall take you to the doorway to this greater Pluralism.

Before starting on this path, let me fend off an easy confusion between Pluralism in Logic and Pluralism in Mathematics. *Logical* Pluralism[1] is a more established position than is Pluralism in Mathematics. A Logical Pluralist accepts different formal logical systems (usually propositional, modal, first-order or possibly second-order logic). See Haack (1978), Da Costa (2007), Batens (2002), for examples of logical Pluralists. See Beall and Restall (2006) for a fully articulated position called 'Logical Pluralism'.

In contrast, Pluralism in the philosophy of mathematics is a Pluralism about different *mathematical theories* as developed and practiced by recognised mathematicians.[2] What Pluralism in Theories amounts to is (at least) Pluralism in: epistemology, ontology, methodology and truth. There is no well worked-out philosophical position called Pluralism in Mathematics in print, to date. Sketches of positions can be found in Aberdein (2005), Bueno (2011) and Friend (2012), amongst others. Moreover, many present-day philosophers of mathematics work in a very pluralist way, but since there is no properly worked-out philosophical position, they are rarely specific about the delimitations of their Pluralism. Many philosophers, computer scientists, logicians and mathematicians tell me that they are pluralist, and, when they say that, they mean quite different things

---

[1] I am using capitals to name a philosophical position, as opposed to naming an attribute of a philosophical position. For example, "pluralism is an attribute of some forms of Structuralism", but "Pluralism is a philosophical position".

[2] Here, 'recognition' is determined socially and institutionally. This is not to say that one has to be a university professor, in order to be considered to be a mathematician. Rather, one has to be engaged in the practice of mathematics, as it is recognised by said institutions. I think that this is the best the Pluralist philosopher can do (since he, or she, recognises no unique foundation for mathematics). The socio-institutional criterion is imperfect by its own standards, since, in retrospect, someone rejected by the institution (in the wide political sense) might later be recognised as an important contributor to mathematics. Similarly, someone judged as 'having an impact' might later fade into the background, as we discover better techniques or favour other theories.

by the term. For lack of a well worked-out, and in print, exposition of Pluralism as a philosophy of mathematics, this paper too is a sketch.

Nevertheless, I can say a few things to help us understand Pluralism better. Pluralism in mathematics is most starkly contrasted to the more traditional philosophies of mathematics which contributed to, and were developed in reaction to, the crisis in the foundations of mathematics. In these traditional philosophies, we propose one mathematical theory as a foundation for the whole of mathematics. With philosophical maturity, the founding theory becomes a norm for mathematical practice, or is intended to set a norm for future practice.[3] Examples include Realism, Logicism, and some versions of Constructivism, Naturalism, Formalism, Fictionalism and Structuralism. In contrast, Pluralism in Foundations (which is the first position we shall visit) rejects the idea of one foundation (in particular, in the form of a formal mathematical theory). The Pluralist in Foundations also rejects the normativity, nay prescriptivity, which accompanies traditional theories. Judgments about mathematical theories or theorems or hypotheses etc. are made from within a context, which is revisable, one standpoint amongst others (from which to make a judgment), and, therefore, not absolute. This is a lesson we can learn from the crisis, and it has taken us a while to learn it. The lesson was not learned in one step. Between our extreme Pluralism and the monism of the traditional positions, we can find some 'in-between' philosophies.

Philosophers of mathematics, who break from the more traditional positions are characterised, here, as favourably using the term 'pluralism' to name a characteristic of their position. We find the term used in some versions of: Structuralism, Formalism, Fictionalism or Naturalism.

In contrast to these in-between positions, the Pluralist places 'pluralism' as the chief virtue of his version of Pluralism. The Pluralist is inspired by the in-between positions, but goes beyond them. Forthwith, I restrict attention to the more traditional positions and

---

[3] The founding theory is a norm in the following sense: purported mathematical work is not recognised as proper mathematics if it cannot, in principle, be reduced to the founding theory. There are strong philosophical objections to foundational theories. See, for example, (Vopeňka 1979), (Bueno 2011).

Pluralism, ignoring the in-between positions.[4] Traditional positions support one foundation to mathematics. The path to Pluralism starts by acknowledging several.

1. Preliminaries: Organisation, Argument and Approach

In this paper, I shall not give a strong direct argument for Pluralism. Instead, I shall discuss some motivating considerations, and then make an indirect argument, by giving a sense of the philosophical work which is done by a 'Pluralist' who takes on one version of Pluralism. To re-emphasise: I write 'one version' because there are many versions. The Pluralist is pluralist about Pluralism! The versions vary along at least two axes. One is: what it is one is pluralist about (this is not the same as ontology – it has to do with subject matter); the other is propositional: how one deals with the issues of truth and the conflict between claims, in and about mathematics.[5]

> *The Pluralism I shall start with (but revise) is a Pluralism about foundations in mathematics, the defender of which, insists that truth can only make sense within a mathematical theory.*

For such a Pluralist, a mathematical theory is not true in itself. It can only be true with reference to a meta-theory (usually by showing it to be equivalent to, or reducible to, that meta-theory). Such is familiar from Shapiro's Structuralism. This Pluralist arranges discussions in and about mathematics into a hierarchy of theories and meta-theories, distinguishing claims made *within* a theory from claims made *about* a theory, from the standpoint of a meta-theory.

The reason I start with this version of Pluralism is to meet a traditional philosopher of mathematics halfway. I can start with philosophical positions which are recognisable to her. As we shall see,

[4] For detailed comparisons of the in-between positions and Pluralism, see (Friend 2012).

[5] Strictly speaking, the default Pluralist view is that 'truth' is handled by a formal logical theory, or at least, the formal theory is a guide to how we are to deal with questions of truth. Are the logical theories, themselves true or false? No. Similarly, the mathematical theories are neither true nor false.

Pluralism in Foundations is unstable. We shall modify it in due course.

I anticipate various reactions. If the presented Pluralist positions strike the reader as anathema to the philosophy of mathematics as it should be practiced, then more detailed, careful and convincing arguments are called for. But this is not the place for those. If the positions resonate with the reader, he might already be a Pluralist. If the positions and example of work done by the Pluralist appeal to the reader, he might be interested in looking further into Pluralism as a philosophy.

2. The Pluralist about 'Foundations' of Mathematics

We begin with a simplifying assumption, in order to present the initial position. The assumption will be revised when we set the Pluralist to work. The simplifying assumption is:

> *a proposed foundation (to mathematics) is a mathematical theory, and mathematical theories are individuated by a language, a set of axioms and some inference rules.*

Under this assumption, we can list different purported 'foundations'. Some foundations of mathematics include Zermelo-Fraenkel set theory with an axiom of choice and consistent extensions of ZFC. If we take the simplifying assumption seriously, then we can also count as foundations: non-well-founded set theories, type-theories, category theories and any other theory to which a lot of mathematics can, in principle, be reduced.[6] Finding such a foundation is the goal of the more traditional philosophies of

[6] If we do not take the assumption seriously, then we might think that all of these theories are much on a par, and are all aspects, or parts, of 'one big founding idea' of mathematics. Woolley language aside, in this case, the word 'theory' is used in a much looser sense than what is proposed in the simplifying assumption. To keep the discussion focused, we make the simplifying assumption, and we can look at a looser idea of theory or foundation later.

mathematics. In contrast, the Pluralist accepts that there are several big theories – many of which are plausible.[7]

> *The Pluralist in Foundations is agnostic as to which is the correct foundation, and over whether there is one unique foundation.*

Note that when we speak and write about several 'foundational theories' co-existing, the very term 'foundation' changes meaning. The building metaphor no longer works. It makes no sense to have several foundations for one building called 'mathematics'. Since the Pluralist is agnostic as to whether there are one or several 'foundations', and in order not to beg any questions, we speak of 'umbrella' theories.

3. Considerations Supporting the Pluralist's Agnosticism

With the shift in vocabulary from 'foundational theory' to 'umbrella theory', we draw attention to the distinction between our philosophical *aspirations* and the philosophical claims we can defend in argument. Pluralists might *hope* that mathematics will one day turn out to be unified in one umbrella theory, but, and this is important, such hope is simply a subjective private feeling, and is not supported on present evidence. So, similarly, a Pluralist might hope (maybe perversely) that there are several irreducible umbrella theories in mathematics, and that there will never, nor can ever, even in principle, be a way of unifying these. Again this is a hope and a private conviction. The Pluralist, as a (public) philosopher is simply agnostic on the issue of unification of mathematics; *mutatis mutandis* for the idea of 'truth of' a theory. No mathematical theory is true *tout court*. A mathematical theory can only be true relative to another theory (by being shown to be reducible to (or embeddable in) that theory). Whether the reducing theory is true, will depend again on another reducing theory. In fact, for the Pluralist, trying to determine the truth of a theory, in the absence of a meta-theory, or reducing theory is labour lost. It is more informative to frankly discuss embeddings, reductions, equi-consistency proofs and other limitative

[7] Keep track of the word 'plausible'. We shall revisit it.

results on a mathematical theory than the truth of the theory.[8] This agnostic attitude makes sense of the following remark, which I quote because I find it representative of some mathematicians' philosophical discomfort. Discussing the truth and independence of the continuum hypothesis, they write:

> *These logical results [of the independence of CH] do not settle the question originally asked by Cantor whether CH or GCH are true or false statements. However, it must be said that these seemingly obvious questions are not very clear: the concepts of truth and falsity (as opposed to the concept of the derivability from axioms) do not have a clear meaning in abstract set theory. Thus we cannot rule out the possibility that Cantor's original questions will turn out to be simply meaningless. (Kuratowski and Mostowski 1976, p. 290).*

The Pluralist's agnosticism is motivated by the following considerations:

(1) there are several, non-equivalent, umbrella theories – many of which have natural philosophical alliances (philosophies which accompany that umbrella theory). Call an 'umbrella theory plus its natural philosophy' a 'traditional position'.

(2) No traditional position is accepted by all mathematicians or by all philosophers.

(3) Moreover, there does not seem to be an immanent convergence of opinion (as we can see in a number of the papers in this volume).

---

[8] It is for this reason that 'Pluralism' is favoured as a name over 'Relativism' since the issues of truth and ontology are side-issues for the Pluralist. In contrast, the Relativist is first and foremost concerned with truth and ontology. Pluralism is related to Relativism in the following way: Pluralism is a mature Relativism. The Pluralist goes beyond Relativism because he is relativist along more dimensions than truth and ontology. The more mature Pluralist position is relativist about truth, ontology, mathematical theories, epistemology, logic and methodology.

Depending on how finely we want to distinguish traditional positions, we might even think that there are a large number of candidate potential positions not yet developed or explored.[9] It seems then, that rather than look for one foundation for mathematics, we should learn to accept the idea that there are several umbrella theories, several potential traditional positions, many of which are plausible; and therefore, no traditional position is correct in stating that their championed big theory represents 'mathematics'. This motivates a further modification of language. Rather than speak of *the true* mathematical theory (or foundation), we can speak of '*plausible* umbrella theories'.

When we replace 'truth of a theory' with 'plausibility of a theory' a few interesting things happen. First note that 'plausibility' is a relative term. Nevertheless, positions are rarely *equally* plausible. Sometimes, one position is more plausible than another on grounds of philosophy of logic. For example, trivial theories are implausible. We could also try to argue that one mathematical theory is implausible because it is illogical or unreasonable. The strongest way to make such an argument is by appeal to a particular formal logic. For example, we might decide that a mathematical theory is implausible because it endorses unrestricted choice, which is not allowed in a constructive logic. However, in the light of a Pluralism about formal logical systems, such an argument begs the question. Of course, not all of us are Logical Pluralists, but, today, logical monists have to defend their monism as well as their choice of logic. We could try to appeal to an informal notion of reasoning and logic, but this is usually unsatisfactory, since ambiguous (i.e., underdetermined) and therefore leading to disagreement (only solved by appeal to a formal system of logic). So the only sure lesson we can learn by appeal to logic or reasoning is to dismiss trivial theories as implausible. But this will not be very helpful in narrowing the field of plausible foundations!

Thankfully, the contrast between trivial and non-trivial mathematical theories is not the only recourse we have to rate the plausibility of a mathematical theory. Plausibility judgments can also

---

[9] We have to be careful about judgments about 'large numbers of positions', especially, if we consider not-in-print positions. First we have to individuate positions, next we have to decide what the parameters are on possible positions.

depend on background knowledge. Presumably, a budding mathematician with little experience and a very narrow area of specialisation is in a lesser position to make a judgment about plausibility, since what is plausible to her will depend on her experience, and resemblance with what she knows of mathematics. In contrast, a mathematician with a vast experience of mathematics and mathematical theories, might well decide that one umbrella theory is more plausible than another. An example is Gödel, although, strictly speaking, he would not have accepted our simplifying assumption. So, knowledge and experience can be used to partially evaluate plausibility. For this reason, attributing plausibility partly comes from an educated judgment, not a mere declaration of taste.

Alas, even amongst the most educated, we do not have consensus, nor do we seem to be heading towards a convergence of judgments. Future arguments and future information might lead to convergence, divergence, or convergence followed by divergence, divergence followed by convergence; we simply do not know. Under these considerations, we have no rational basis, nor authoritative basis (based on amount of education),[10] nor inductive basis, upon which to make a choice for one position as the true position.[11] It is for this reason that the Pluralist demurs from making a choice, and accepts all plausible umbrella theories as on a par, in the first instance (i.e., they will be subjected to further evaluation since we are aware that 'plausibility' is a relative term requiring qualification). The Pluralist is a principled agnostic (as opposed to a lazy agnostic).[12] For the

[10] 'Amount of education' is, of course, not to be confused with number of degrees or prestige of award granting institutions. Here 'education' is meant in the basic sense of pursued, sustained and critical enquiry. University degrees are a rough indicator.

[11] We can work to find the most plausible big theory, but given that to determine what is the most plausible theory, we have to make some choices which will seem arbitrary or questionable to some (respectable people), the purpose of so doing becomes unclear; unless we take it as a private project responding to our private convictions. To remain in the public sphere, we are rationally more secure in noting different notions of plausibility, and evaluating which theories fit which notions.

[12] A lazy agnostic is not interested in the debate (any more). So she gives up, and just says she is agnostic. A principled agnostic is interested in comparing, evaluating the comparison, changing his mind and being honest

principled agnostic, the traditional philosophies are not supportable on present evidence.

Some readers will be bristling, and I leave them in this state. Instead of giving a knock-down defence of Pluralism, or an exposition of the whole Pluralist position, I turn to what it is that the Pluralist does, as a philosopher of mathematics. The purpose is to fend off the *reductio* argument against the Pluralist that the Pluralist cannot practice philosophy. Rather, what is left to her is: sociology, psychology, history, historiography and so on – anything but philosophy. Actual, therefore possible, as the Mediaevals taught us. Since Pluralism in Foundations is unstable, I, instead, display the practice of the Pluralist in Mathematical Theories who is interested in the question of foundations. Details about his position, or rather, the family of positions he could occupy, need not concern us here.

4. Setting the Pluralist in Mathematical Theories to Work

To stay with the same theme of 'foundational mathematics', the Pluralist does not ignore and merely dismiss the foundational aspiration. *Au contraire*, the Pluralist concedes that there is something to the claim that ZFC is a 'foundation', even if we have learned to call it an 'umbrella' theory. After all, a number of very well respected mathematicians take ZFC, or something like it, to be very important. The observation is not lost on the Pluralist, who will now work with this observation. We can start by revising our simplifying assumption, and say something more accurate, namely that:

> *set theory (not ZFC, which is one formal representation of set theory, but whatever it is that ZFC+ is meant to represent) is really used as a 'lingua franca' or a 'reference point' by mathematicians (usually by appeal to the formal theory).*

---

and industrious towards the question. He is also well aware that there might be no (definitive or even temporary) resolution to the issue under question. Maybe 'industrious agnostic' would be a better term. 'Principled' is meant to refer to the idea that there is a defence the agnostic can give for his agnosticism, and that he does care deeply about the outcome of the debate, but more than this, he cares about the honesty and strength of argument used in the debate.

We have modified our initial simplifying assumption in two respects. One is that we are not individuating theories by formally presented theories. The other is that we are talking of reference points rather than foundations. In some mathematical circles, set theory is 'central' to mathematics. Since the Pluralist is agnostic about the foundational claim, he wants to know what it is that is appealing in the claim that ZFC is a foundation, or why it is that set theory is a *lingua franca*, or a reference point for mathematicians. The claim about ZFC is descriptive. Because of this, answers can be sought by: sociologists, historians, historiographers, psychologists and so on. But the Pluralist can also ask a more philosophical question: "Is set theory really the *best* reference point for an arbitrary mathematical theory?" Or, similarly: "Should we continue to use ZFC as a *lingua franca*?" Both are hopelessly vague and ambiguous questions because of the words 'best' and 'should'. The Pluralist wants to give a rigorous and defendable answer, and we can do this. Let us begin by making some of the terms more precise: 'set theory' (most loyally represented by which formal theory?), 'the best' (according to what measure?), 'reference point' (what is the difference between reference point and foundation?), 'arbitrary mathematical theory' (what gets counted?). But we want to do more than this. We want to bear in mind that when we make these terms more precise, we set parameters on the question, and Pluralists are aware that these parameters can be revised. With that caution in mind, let us get to work and propose disambiguations of the assumption.

Working backwards, an 'arbitrary mathematical theory' is a theory, the inventor, discoverer or developer of which, believes that it is a mathematical theory. It has a formal language, which expresses all of the truths, or theorems, in the theory. It might be presented as a mathematical structure (as defined by model theory), as an axiomatic theory, as a rule-based theory, and possibly in some other way. But we cannot allow anything we wish. To give a precise answer to our question, within revisable but temporarily set parameters, we need 'arbitrary theory' to be comparable to set theory. To allow comparability, an arbitrary theory is either (i) a formally presented theory, in which case, let us modify 'set theory' to mean formal ZFC, or some appropriate version: ZF, ZFC+ etc. Or, (ii) our arbitrary theory is not formally presented (or is highly ambiguous between different formal presentations) so our comparison will be made with something underlying formally presented set theory, i.e., *that to which* ZFC+ is responsible (or, *against which* attempts at formal representation are judged).

'Reference point' is ambiguous between at least: 'can be reduced to ZFC', 'can be shown to be equi-consistent with ZFC', 'ZFC can be reduced to our theory' and 'shares some recognisable features with ZFC'. The last is the most nebulous. An example is the theory of semisets. The 'recognisable features' include sets and all of the axioms of set theory (some are a little modified), but our theory also includes the theory of semisets, which concerns something less definite than a set, that is, less definite in its membership (set members we can work with are fixed with respect to a perspective, as we change perspective, so the members we can manipulate and work with change). ZF is not entirely reducible to the theory of semisets, but there is a strong basis for comparison.

'The best' as it is used in the context of our question and given our agnosticism, is determined by appeal to the practice of mathematics. It means something of the form 'is recognised by, and is respected by, (a good enough number of) recognised and respected mathematicians today'. This is far from an absolute and pure judgment! Rather, it combines a statistic determined by some sense of the 'institution' of mathematical research and evaluation. This proposed disambiguation of 'the best' makes explicit the instability of a positive or negative answer to the questions, since it is the 'institution of mathematical research' which will be adding the normative force of 'should' or determining what counts as 'best'. Not only does the institution change over time, with the people and their attitudes, but, there will be different ways of working out what mathematicians consider to be the best reference point, depending on what we take to represent the 'institution'.

The last term is 'set theory'. 'Set theory' is what underlies the various formal representations.[13] It is a mathematical theory of sets and membership. To make precise, and fix, what 'set theory' means,

[13] Note that I write as though 'set theory' is one thing being formally represented. The only reason to write this way is for simplifying the conclusion. It is easier to keep some goal posts rigid, for a while. They can be moved later.

we give a formal representation, or several formal representations. Moreover, we often have a sense that the formal representation of set theory is not completed. We extend formal set theories by adding new axioms, and experimenting with these to see what follows when we add them, and try to determine whether or not they are 'fruitful' or 'natural'. Again, the latter are also hopelessly vague and ambiguous terms, which we can try to make more precise, maybe at the expense of our intuitions, in favour of an artificial precision.

5. Answers

Already in disambiguating, we have learned a lot about our question. The analysis of the very general philosophical question "Is set theory really the best reference point for an arbitrary mathematical theory? Or, should it be such?" gives some idea of the *modus operandi* of a Pluralist philosopher of mathematics. Having made his preliminary analysis, the Pluralist can now give an answer to the question.

The Pluralist will want to make a survey to confirm or disconfirm the following guesses.

(1) The first guess is that logicians (broadly construed to include set theorists, model theorists, logicians working in different traditions for propositional, first-order and second-order logics, proof theorists, some computer scientists, and maybe some others) tend to feel *more secure* if they can show that their theory is reducible to ZFC, or some formal theory which has ZFC as a core of the theory. They feel *secure* if their theory is equi-consistent with ZFC. That is, ZFC, as a reference point affords a sense of security for logicians.

(2) It is the *de facto practice* amongst logicians to use ZFC as a default reference point.

(3) Other logicians turn the tables. They *use* 'set theory' in the sense of 'what underlies the formal representation' as a core for recognising, developing and interpreting formal umbrella theories. They treat formal set theory as a starting point, not as a foundation. For example, set theory is what is used to compare what happens

when we make changes to ZFC, by adding new axioms. So 'set theory' in this sense, is assumed as a reference point *ab initio*. These logicians are set theorists in the broad sense of working in and around ZFC, and seeing mathematical theories in terms of ZFC.[14]

Indicators that these guesses are correct include the use of the notation we find in set theory textbooks, and the use of language, axioms and theorems familiar from set theory in the development of their own theory. These guesses concern logicians. What about other mathematicians? The speculation continues:

(4) some (respected *mathematicians*) *do not* use set theory as a reference point, and some *cannot*.

If the fourth guess is correct, then, I speculate that this is because amongst 'real working mathematicians' (*pace* the traditional philosopher of mathematics) very little attention is paid to set theory, or ZFC. They are usually not trained in set theory, and they are untroubled by foundational issues. They are also not concerned with finding 'a unique reference point', or default reference point for the theory (or theories) they are working in. Their sense of security is based in practice. It is enough if those judging, evaluating and correcting their work endorse it, and these endorsers will be people in the same or related fields of work in mathematics.

Is this a question of 'distribution of labour' (so all mathematical work is traceable, in some way, by more specialised mathematicians to set theory) or is it really that set theory simply does not act as a reference point? To answer this question, imagine if explicit traces of set theory were to disappear. All books about set theory and people working in set theory or with a good working knowledge or concern about set theory were to be shipped to another part of the galaxy, and be *incommunicado* with the mathematicians left behind. In this case I conjecture that there would be little pause in mathematical work, and there would be no urge to re-invent set theory, or find an overarching reference point or foundational theory. I make the conjecture based

[14] I should like to thank Brian Skyrms for pointing out this approach to me in conversation.

on some areas of mathematics where it makes no sense for ZFC to be the reference point for the claims of the theory, such as statistics or calculus.[15] Testing the conjecture is delicate future work.[16]

If my statistical guesses are confirmed by a survey, and my guesses are supported by further evidence and argument, then the Pluralist answer will be that:

> *a lot of logicians use ZFC as a reference point, but this is not representative of all mathematicians.*

6. Conclusion

Having gone through our careful exercise, we will also understand a number of subtleties attending the answer, and we can confirm or disconfirm and revise my guesses. We also see that we have to leave behind the territory familiar to the traditional philosopher of mathematics – leave ZFC, and look to the actual practices in mathematics.

More broadly, other Pluralist work includes: showing more precisely how Pluralism is different from the in-between positions. Recall that these are some versions of Structuralism, Formalism, Fictionalism and Naturalism. Most of these comparisons can be found elsewhere.

---

[15] We can, of course, reduce, or interpret calculus or statistics in terms of set theory, but being able to, and having to do so, or needing to do so are quite distinct. Arguably, reducing calculus or statistics to set theory considerably distorts these areas of mathematics, and is quite unhelpful. This is not to say that a proposed reduction is completely unhelpful and uninformative. Rather, it is considerably inefficient if one is interested in calculus – at least on present evidence. To insist otherwise is to run the risk either of begging the question (so mathematics is just defined in terms of reference to set theory), or to make the term 'set theory' so broad as to be vacuous: so under the imagined scenario of losing trace of set theory, all of mathematics and mathematicians would be shipped. As I write, the issue is delicate, and the strength of conjecture rests on such work.

[16] We would have to navigate between Scylla and Charybdis. Scylla: if we take the term 'explicit mention' broadly, we will find that all, or almost all, mathematics (post 1870s) would disappear. Charybdis: we might think that set theory would still be 'there' underlying most mathematical thought.

Nevertheless, with the constrained exercise we performed here, we have indicated a way to embrace the crisis in the foundations of mathematics. We can now stand in the doorway to other forms of Pluralism in mathematics.

Bibliography

ABERDEIN, A., 2005. *Hybrid Pluralism*. http://arxiv.org/abs/math/0505034.

BATENS, D., 2002. 'A General Characterisation of Adaptive Logics', Manuscript.

BEALL, J.C. and RESTALL, G., 2006. *Logical Pluralism*. Oxford: Clarendon Press.

BUENO, O., 2011. Relativism in Set Theory and Mathematics. *A Companion to Relativism.* (Ed.) S.D. Hales. Oxford: Wiley-Blackwell, pp. 553-568.

COLYVAN, M., 2001. *The Indispensability of Mathematics.* Oxford: Oxford University Press.

DA COSTA, N.C.A., KRAUSE D. and BUENO, O., 2007. Paraconsistent Logics and Paraconsistency. (ed.) D. Jacquette. *Philosophy of Logic.* Amsterdam: North Holland, pp. 791-911.

FIELD, H., 1980. *Science Without Numbers.* Princeton: Princeton University Press.

FRIEND, M., 2012. Pluralism and "Bad" Mathematical Theories: Challenging our Prejudices. *Paraconsistency: Logic and Applications*. (Eds.) Koji Tanaka, Franz Berto, Edwin Mares and Paoli Francesco. Springer.

HAACK, S., 1978. *Philosophy of Logics.* Cambridge: Cambridge University Press.

HELLMAN, G., 1989. *Mathematics Without Numbers*. Oxford: Oxford University Press.

KÖRNER, S., 1962. *The Philosophy of Mathematics; an Introductory Essay.* New York: Harper and Row.

MADDY, P. 1997. *Naturalism in Mathematics.* Oxford: Clarendon Press.

SHAPIRO, S., 1997. *Philosophy of Mathematics; Structure and Ontology.* Oxford: Oxford University Press.

VOPEŇKA, P., 1979. *Mathematics in the Alternative Set Theory*. Leipzig: Teubner –texte zur Mathematik.

La méthode axiomatique durant la crise des fondements

Mathieu Bélanger
Université de Montréal

Résumé

Ce texte examine la méthode axiomatique durant la période de l'histoire des mathématiques correspondant à la crise des fondements. Il a pour objectif de montrer que la méthode axiomatique changea, mais aussi de comprendre la nature de ces changements. À cette fin, les conceptions de Frege, Hilbert et Noether sont analysées.

L'objectif du présent texte est de montrer que la méthode axiomatique changea durant la crise des fondements, c'est-à-dire durant cette période de l'histoire des mathématiques comprise entre 1885 et 1930 et, ce faisant, de mettre en évidence la nature de cette évolution.

Il importe de préciser qu'il n'a pas la prétention de faire une histoire exhaustive de la méthode axiomatique durant la crise des fondements, mais plutôt de se concentrer sur ce qui apparaît être des moments charnières. À cet égard, trois moments pouvant être associés à autant de mathématiciens ressortent : Gottlob Frege, David Hilbert et Emmy Noether.

L'analyse proposée se divise en quatre sections. Il sera dans un premier temps question de la méthode axiomatique chez Frege. La deuxième section se penchera quant à elle sur les *Grundlagen der Geometrie* de manière à comprendre la conception de Hilbert. Afin de bien mettre en lumière la rupture que marqua Hilbert, la suivante reviendra sur la correspondance de Frege et Hilbert en mettant l'accent sur le rôle des définitions et la vérité des axiomes. Finalement, dans la dernière section, l'utilisation novatrice que fit Noether de la méthode axiomatique sera analysée par le biais de ses travaux sur le théorème de factorisation unique.

1. Frege : les axiomes comme vérités évidentes

Dans *Die Grundlagen der Arithmetik* (Frege 1884), Frege formula la thèse selon laquelle les vérités mathématiques sont analytiques. Ce faisant, Frege s'opposait principalement à Kant selon qui les propositions arithmétiques sont synthétiques *a priori*, mais aussi à Mill qui les considérait *a posteriori*.

Avant d'aller plus loin, il semble essentiel de rappeler que, selon Frege, les notions de propositions analytiques, synthétiques, *a priori* et *a posteriori* dépendent du processus de justification.

> Now these distinctions between a priori and a posteriori, synthetic and analytic, concern, as I see it, not the content of the judgement but the justification for making the judgement. Where there is no such justification, the possibility of drawing the distinctions vanishes. (…) When a proposition is called a posteriori or analytic in my sense, this is not a judgement about the conditions, psychological, physiological and physical, which have made it possible to form the content of the proposition in our consciousness; nor it is a judgement about the way in which some other man has come, perhaps erroneously, to believe it true; rather, it is a judgement about the ultimate ground upon which rests the justification for holding it to be true. (Frege 1884, 3)

Ainsi, déterminer le caractère analytique, synthétique, *a priori* ou *a posteriori* d'une proposition donnée revient à examiner les étapes de sa démonstration, depuis la conclusion finale jusqu'aux « vérités primitives », pour reprendre l'expression de Frege (1884, 4).

Ceci conduisit Frege (1884, 4) à adopter les définitions suivantes. Une proposition est *analytique* si sa démonstration ne repose que sur les lois de la logique et des définitions. En contrepartie, elle est *synthétique* si sa démonstration emploie des vérités qui appartiennent à une science particulière, c'est-à-dire au moins une proposition qui ne relève pas de la logique. Par ailleurs, une proposition est *a priori* si sa démonstration ne dépend que de lois générales qui n'ont elles-mêmes besoin d'aucune preuve. Elle est plutôt *a posteriori* s'il est impossible d'en construire une démonstration sans recourir à des vérités qui réfèrent à des objets particuliers.

Pour établir l'analycité des propositions arithmétiques, Frege devait donc montrer qu'elles pouvaient toutes être démontrées à partir des lois de la logique et de définitions

Frege s'attela à cette tâche dans *Grundgesetze der Arithmetik* (Frege 1964). Dans un premier temps, il formula des axiomes exprimant des vérités logiques – les célèbres lois de base –, des définitions donnant une dénotation aux concepts intervenant dans ces axiomes de même que des règles d'inférence[1]. Dans un deuxième temps, il démontra certaines propositions arithmétiques à partir des lois de base en utilisant seulement les règles d'inférence. À ce propos, si Frege croyait sa réduction de l'arithmétique à la logique sans faille – il écrit dans l'introduction que « … the only criticisms that can justly be made against the book concern not the rigor but merely the choice of the course of proof and of intermediate steps. » (1964, vi) –, le paradoxe de Russell le força, pour employer un euphémisme, à reconsidérer le succès de son projet.

Afin de bien comprendre l'évolution subséquente de la méthode axiomatique, il importe d'examiner les notions de définition et d'axiome telles que les conçoit Frege.

Premièrement, le rôle des définitions est de fournir des noms pour désigner des concepts. Dans cette optique, leur raison d'être tient à ce qu'elles permettent de désigner avec plus de brièveté : « The definitions do not really create anything, and in my opinion may not do so; they merely introduce abbreviated notations (names), which could be dispensed with were it not that lengthiness would then make for insuperable external difficulties. » (Frege 1964, vi) Qui plus est, Frege s'insurge contre l'idée qu'une définition permette d'introduire un nouveau concept ; un nom ne peut être attribué qu'à des objets dont l'existence est établie.

[1] Pour un résumé des lois de base et des règles d'inférence, voir Frege (1964, §47-48).

> It is important that we make clear at this point what definition is and what can be attained by means of it. It seems frequently credited with a creative power; but all it accomplishes is that something is marked out in sharp relief and designated by a name. Just as the geographer does not create a sea when he draws boundary lines and says: the part of the ocean's surface bounded by these lines I am going to call the Yellow Sea, so too the mathematician cannot really create anything by his defining. Nor can one by pure definition magically conjure into a thing a property that in fact it does not possess – save that of being called by the name with which one has named it. (1964, xiii)

Deuxièmement, selon Frege, les axiomes, c'est-à-dire les lois de base, sont des vérités évidentes[2]. Ceci signifie qu'ils ne requièrent aucune justification. Contrairement aux autres propositions, les axiomes sont vrais en eux-mêmes au sens où leur vérité ne dépend d'aucune autre proposition. C'est en vertu de ce statut épistémologique que les axiomes peuvent être le fondement de la vérité d'autres propositions. Ceux-ci étant vrais, toute proposition qui en sera déduite par l'entremise de règles d'inférence valides sera elle-même vraie.

Les axiomes ont donc un rôle fondationnel chez Frege. En effet, les axiomes constituent le fondement de la connaissance arithmétique dans la mesure où la vérité de toute proposition arithmétique dépend de la vérité des axiomes. Ces derniers constituent en ce sens la justification ultime des propositions arithmétiques.

2. Hilbert : la méthode axiomatique comme méthode de clarification

L'objectif de Hilbert dans *Grundlagen der Geometrie* (1956, 1971), d'abord publié en 1899, était l'identification d'un système d'axiomes minimal pour la géométrie, c'est-à-dire qu'il cherchait à formuler les axiomes les plus simples possibles à partir desquels toute proposition

[2] Pour une analyse approfondie de la notion d'évidence chez Frege, voir Jeshion (2001).

géométrique pourrait être déduite, mais aussi à résoudre la question de l'indépendance des axiomes de la géométrie, en particulier les axiomes des parallèles et d'Archimède. (Hilbert 1971, 10; Frege 1980, lettre 4, 38)

Hilbert prend pour point de départ trois types d'objets qu'il nomme respectivement « points », « droites » et « plans ». À ce stade-ci, « point », « droite » et « plan » ne sont que des termes pour désigner les objets de chaque type.

Ce que sont les points, les droites et les plans est entièrement défini par les relations qui existent entre eux et qui sont exprimées par les termes « être sur », « entre », « parallèle », « congruent » et « continu ». Ces relations sont quant à elles entièrement décrites à l'aide d'axiomes : les axiomes de la géométrie.

Hilbert organise les axiomes de la géométrie en cinq groupes qui décrivent chacun une des relations à l'œuvre en géométrie :

> 1- Axiomes d'appartenance : les axiomes d'appartenance définissent les relations d'appartenance et d'inclusion entre les points, les droites et les plans.
>
> **Exemple**. Pour toute paire de points distincts A et B, il existe une droite a passant par ces points.
>
> 2- Axiomes d'ordre : ces axiomes définissent la relation « entre », ce qui permet de déterminer l'ordre des points dans l'espace.
>
> **Exemple.** Pour toute paire de points A et C, il existe au moins un point B qui appartient à la droite AC et qui soit entre A et C.
>
> 3- Axiomes de congruence : comme le nom du groupe le suggère, ces axiomes définissent la notion de congruence.
>
> **Exemple.** Si A et B sont deux points d'une droite a et A' est un point de a ou d'une autre droite a', alors, il est toujours possible, d'un côté donné de A' de trouver un point B' tel que

le segment AB soit congruent ou égal au segment $A'B'$, ce qui s'écrit $AB \equiv A'B'$.

4- Axiome des parallèles : ce groupe ne contient en fait qu'un seul axiome, à savoir l'axiome des parallèles, lequel définit, comme son nom l'indique, la notion de parallélisme.

Exemple. Si a est une droite et A un point qui n'appartient pas à a, alors, dans le plan déterminé par a et A, il existe au plus une droite qui passe par A et qui ne coupe a, c'est-à-dire qui est parallèle à a.

5- Axiomes de continuité : ce groupe contient l'axiome d'Archimède qui permet d'introduire la continuité et l'axiome de complétude qui affirme la complétude du système.

Exemple. Si AB et CD sont deux segments quelconques, alors il existe un nombre entier n tel que le segment CD répété n fois à partir du point A sur la demi-droite déterminée par B mène à un point situé au-delà de B.

D'après Hilbert, ces axiomes ont pour source l'intuition, l'intuition spatiale pour être plus précis. À cet égard, la formulation des cinq groupes d'axiomes est précédée de l'affirmation suivante : « chacun de ces groupes exprime quelques faits fondamentaux, liés les uns aux autres et qui nous sont donnés par l'intuition. » (1971, 11) Les axiomes et leur organisation en ces cinq groupes expriment et systématisent certaines intuitions de base à propos de l'espace. Il faut toutefois insister sur le fait que Hilbert ne dit pas que ces axiomes sont vrais en soi. Contrairement à Frege, bien qu'ils aient pour source l'intuition, les axiomes de la géométrie ne sont pas des vérités évidentes.

Sur la seule base de ces axiomes, Hilbert est en mesure de récupérer toutes les propositions de la géométrie euclidienne. L'omission de certains groupes d'axiomes lui permet également d'obtenir certains théorèmes valides dans des géométries non euclidiennes. De plus, Hilbert obtient des résultats qui relèvent de la géométrie projective en ce qu'il démontre, dans les cinquième et sixième chapitres respectivement, les théorèmes de Desargues et de Pascal sans recourir aux axiomes de continuité.

Dans le deuxième chapitre, Hilbert s'attaque à la cohérence et à l'indépendance des axiomes de la géométrie. Il construit d'abord un corps de nombres algébriques et montre à l'aide de celui-ci la cohérence relative des axiomes de la géométrie et de l'arithmétique des nombres réels. Ceci signifie que la cohérence des axiomes géométriques est équivalente à celle de l'arithmétique des nombres réels et donc que toute contradiction dans l'une se manifesterait dans l'autre.

Dans un deuxième temps, Hilbert démontre que les groupes d'axiomes, mais aussi que les axiomes d'un même groupe sont indépendants. L'approche de Hilbert consiste à construire, pour chaque axiome A, un modèle d'une théorie géométrique qui satisfait tous les axiomes sauf A.

Comme le souligne Corry (1996), dans *Grundlagen der Geometrie*, Hilbert utilise donc la méthode axiomatique pour dévoiler la structure logique de la géométrie et, ce faisant, clarifier cette théorie mathématique. Premièrement, la définition des concepts géométriques – c'est-à-dire ce qu'est un point, une droite, un plan, etc. – dérive de la structure logique de la théorie. Les mots « point », « droite », « plan », etc. n'ont un sens qu'en vertu de leur participation à cette structure. Deuxièmement, l'approche axiomatique de Hilbert met en lumière les relations de dépendance entre les propositions géométriques et les divers axiomes des cinq groupes. En d'autres termes, les démonstrations permettent de comprendre de quels axiomes une proposition donnée dépend. Finalement, l'analyse des propriétés logiques des axiomes de la géométrie contribue à la connaissance des propriétés métalogiques des diverses théories géométriques. Par exemple, Hilbert (1971, 57) affirme que la cohérence de la géométrie non euclidienne est une conséquence de la cohérence de la géométrie euclidienne.

Par le fait même, Hilbert ne voit pas la méthode axiomatique comme un outil de création de nouvelles théories mathématiques par la simple formulation d'axiomes arbitraires. Corry écrit à ce propos :

> Hilbert's own conception of axiomatics did not convey or encourage the formulation of abstract axiomatic systems:

> his work was instead directly motivated by the need to better define and understand existing mathematical and scientific theories. In Hilbert's view, the definition of systems of abstract axioms and their analysis following the above described guidelines, were meant to be conducted for established and elaborated mathematical entities. Hilbert saw natural science – and in particular geometry within it – as an organic entity, growing simultaneously in various directions. In this view, the development of science involved both an expansion in scope and an ongoing clarification of the logical structure of its existing parts. (1996, 162)

Chez Hilbert, les axiomes jouent donc un rôle de clarification de théories mathématiques déjà existantes. En cela, sa conception s'opposait à celle de Frege qui, tel que vu précédemment, attribuait plutôt aux axiomes un rôle fondationnel. Or, ce point de divergence n'était pas le seul.

3. La correspondance de Frege et Hilbert

Historiquement, *Grundlagen der Geometrie* fut à l'origine d'une brève correspondance entre Frege et Hilbert. Dans celle-ci, Frege s'inscrit en faux avec la conception des axiomes et des définitions mise de l'avant par Hilbert. En fait, ses objections permettent de prendre conscience de la transformation radicale que fit subir Hilbert à la méthode axiomatique et, du coup, du fossé séparant les conceptions frégéenne et hilbertienne. Comme l'écrit l'éditeur dans l'introduction de la section consacrée à la correspondance de Frege et Hilbert :

> Although Frege's logical objections were well taken, and although a correct understanding of the axiomatic method must begin with Frege, the dominant view, especially among mathematicians, is still the view expressed recently by H. Scholz: '… no one doubts nowadays that while Frege himself created much that was radically new on the basis of the classical conception of science, he was no longer able to grasp Hilbert's radical transformation of this conception of science, with the result that his critical

> remarks, though very acute in themselves and still worth reading today, must nevertheless be regarded as essentially beside the point.' (Gabriel et al. 1980, 31)

Dans sa lettre du 27 décembre 1899, Frege reproche à Hilbert de confondre axiomes et définitions. Tel que vu à la section 2, Frege considère que les axiomes expriment des vérités intuitives alors que les définitions ne font que fixer la dénotation des concepts.

> The explanations of sects 1 and 3 are apparently of a very different kind, for here the meanings of the words 'point', 'line', 'between' are not given, but are assumed to be known in advance. At least it seems so. But it is also left unclear what you call a point. One first thinks of points in the sense of Euclidean geometry, a thought reinforced by the proposition that the axioms express fundamental facts of our intuition. But afterwards you think of a pair of numbers as a point (…) Here the axioms are made to carry a burden that belongs to definitions. To me this seems to obliterate the dividing line between definitions and axioms in a dubious manner (…) (Frege 1980, 35).

Du point de vue de Frege, les concepts de « point », « ligne », etc. ne peuvent pas être définis par les axiomes puisqu'un axiome ne peut être formulé s'il fait référence à des concepts qui n'ont pas encore été définis, c'est-à-dire des termes dont la dénotation n'a pas été au préalable fixée. En effet, un axiome exprimant une vérité, comment pourrait-il être vrai s'il réfère à des concepts inconnus ?

> I should like to divide up the totality of mathematical propositions into definitions and all the remaining propositions (axioms, fundamental laws[3], theorems). Every definition contains a sign (an expression, a word) which had no meaning before and which is first given a meaning by the definition. Once this has been done the definition can be turned into a self-evident proposition which can be

[3] Étrangement, Frege ne semble guère faire de différence entre un axiome et une loi fondamentale dans *Grundgesetze der Arithmetik*. Cf. Frege (1964, 3).

> used like an axiom. But we must not lose sight of the fact that a definition does not assert anything but lays down something. Thus we must never present as a definition something that is in need of proof or of some other confirmation of its truth. (…) It is very essential for the strictness of mathematical investigations that the difference between definitions and all other propositions be observed in all strictness. The other propositions (axioms, fundamemtal laws and theorems) must not contain a word or sign whose sense and meaning, or whose contribution to the expression of a thought, was not already completely laid down, so that there is no doubt about the sense of the proposition and the thought it expresses. The only question can be whether this thought is true and what its truth rests on. Thus axioms and theorems can never try to lay down the meaning of a sign or word that occurs in them, but it must already be laid down. (Frege 1980, 36)

Qui plus est, dans une lettre subséquente datée du 16 septembre 1900, Frege affirme qu'il ne peut y avoir de relations entre concepts qu'une fois ceux-ci définis. Bref, définir ce que sont les points, les droites et les plans par le biais des relations qu'ils entretiennent est impossible. Selon Frege, ces relations sont déterminées par ce que les points, droites et plans sont : « There can be talk about relations between concepts – e.g. the subordination of one concept to another – only after these concepts have been given sharp limits, but not while they are being defined. » (1980, 49)

Dans sa réponse du 22 septembre, Hilbert réitère qu'une définition axiomatique est au contraire la seule façon de ne rien présupposer et d'être parfaitement rigoureux, non sans laisser poindre une certaine irritation cette fois...

> In my opinion, a concept can be fixed logically only by its relations to other concepts. These relations, formulated in certain statements, I call axioms, thus arriving at the view that axioms (perhaps together with propositions assigning names to concepts) are the definitions of the concepts. I did not think up this view because I had nothing better to

> do, but I found myself forced into it by the requirements of strictness in logical inference and in the logical construction of a theory. (Frege 1980, 51)

Au-delà du rôle des définitions, Frege et Hilbert ne s'entendent pas sur la question de la vérité des axiomes et de leur cohérence. Toujours dans sa lettre du 27 décembre 1899, l'auteur de *Grundgesetze der Arithmetik* rappelle que les axiomes, en ce qu'ils expriment des vérités évidentes, sont nécessairement vrais et que leur cohérence est une conséquence triviale de ce statut épistémologique.

> I call axioms propositions that are true but are not proved because our knowledge of them flows from a source very different from the logical source, a source which might be called spatial intuition. From the truth of the axioms it follows that they do not contradict one another. There is therefore no need for a further proof. (1980, 37)

Dans sa réponse du 29 décembre, Hilbert maintient au contraire que la cohérence est le critère de vérité des axiomes. Il va même plus loin en faisant de la cohérence des axiomes le critère d'existence des objets mathématiques qu'ils définissent.

> I found it very interesting to read this very sentence in your letter, for as long as I have been thinking, writing and lecturing on these things, I have been saying the exact reverse: if the arbitrary given axioms do not contradict one another with all their consequences, then they are true and the things defined by the axioms exist. This is for me the criterion of truth and existence. (Frege 1980. 39)

Historiquement, si la conception hilbertienne des axiomes s'imposa, la méthode axiomatique elle-même en vint à jouer un nouveau rôle dans le contexte de l'algèbre abstraite au cours des années 1920.

4. Noether : la méthode axiomatique comme outil de travail

Une transformation profonde de la méthode axiomatique survint avec les travaux d'Emmy Noether en algèbre. Cette transformation tient à ce que, chez Noether, la méthode axiomatique est un outil pour le travail mathématique : elle permet de démontrer directement des résultats inédits et non triviaux abstraitement.

Cette utilisation inédite est parfaitement illustrée par les travaux de Noether sur le théorème de factorisation unique. Ce théorème est une extension du théorème fondamental de l'arithmétique qui affirme que tout nombre naturel peut être décomposé de manière unique en un produit de puissances de nombres premiers. Par exemple, $12 = 2^2 \times 3$ et $100 = 2^2 \times 5^2$.

Avant que Noether n'entre en scène, deux types de généralisations du théorème fondamental de l'arithmétique étaient connus.

Premièrement, l'étude des extensions de l'arithmétique élémentaire avait conduit les mathématiciens à vouloir y transposer les propriétés des nombres naturels. Par exemple, Gauss, à qui la démonstration du théorème fondamental de l'arithmétique est habituellement attribuée, démontra un théorème de factorisation pour les entiers complexes[4] alors que Kummer en fit autant pour les entiers cyclotomiques[5]. Ces diverses généralisations avaient toutefois l'inconvénient de ne pas être valides pour certains domaines de nombres. La solution générale vint de la théorie des idéaux telle que formulée par Dedekind. Brièvement, étant donné un corps de nombres, Dedekind montra que tout idéal de l'anneau des entiers algébriques[6] qui appartiennent à ce corps peut être représenté de manière unique comme un multiple de puissances d'idéaux premiers[7].

Deuxièmement, au début du XX^e^ siècle, les concepts et méthodes développés par Dedekind furent transposés à la théorie des polynômes. Une théorie des idéaux de polynômes vit donc le jour et

---

[4] Un *entier complexe* est un nombre de la forme $a + ib$ où $a, b \in \mathbb{N}$.

[5] Un *entier cyclotomique* est un nombre de la forme $a_0 + a_1\theta + a_2\theta^2 + \ldots + a_n\theta^n$ où θ est une racine primitive d'ordre p de l'unité.

[6] Un entier est *algébrique* s'il est la racine d'un polynôme unitaire à coefficients entiers.

[7] Pour une présentation plus détaillée, voir Corry (1996, chapitre 2).

déboucha, principalement grâce aux travaux de Lasker et Macaulay, sur un théorème de factorisation unique pour les polynômes[8].

Au cours des années 1920, Noether unifia toutes les questions de factorisation au sein d'une théorie abstraite. À vrai dire, historiquement, elle proposa deux versions de cette théorie unifiée : la première dans l'article « *Idealtheorie in Ringbereichen* » de 1921 et la seconde, qui peut être considérée comme définitive, dans l'article « *Abstrakter Aufbau der Idealtheorie* » de 1926.

Noether annonce on ne peut plus clairement son projet d'unification dès le début de « *Idealtheorie in Ringbereichen* » : « Le présent travail consiste en une *transposition des théorèmes de factorisation pour les entiers rationnels et pour les idéaux de corps de nombres algébriques aux idéaux d'anneaux d'intégrité arbitraires et d'anneaux généraux*[9]. » (1921, 25)

La pierre de touche de cette unification fut le concept d'anneau. À cet égard, Noether considère des anneaux abstraits, notion qui avait été définie – axiomatiquement – pour la première fois par Fraenkel[10]. En effet, elle vit un lien entre la théorie des anneaux et le théorème de factorisation unique que ses prédécesseurs n'avaient pas vu.

Cela étant dit, à la lumière de l'objectif de la présente section qui n'est pas de comprendre le théorème de factorisation unique dans sa forme générale, seul le premier lemme que démontre Noether dans « *Idealtheorie in Ringbereichen* » sera examiné dans la mesure où il illustre parfaitement son utilisation novatrice de la méthode axiomatique.

Noether commence par définir la notion d'anneau commutatif et précise que sa définition se veut abstraite, c'est-à-dire que la nature des éléments n'est pas spécifiée. Un *anneau commutatif* est un ensemble Σ d'éléments a, b, c,… muni d'une relation d'égalité et de deux opérations, appelées addition et multiplication, qui, à toute paire d'éléments a et b, associent respectivement la somme $a + b$ et le produit $a \cdot b$ telles que les propriétés suivantes sont satisfaites :

[8] À ce sujet, voir Corry (1996, §4.7).

[9] Den Inhalt der vorliegenden Arbeit bildet die *Übertragung der Zerlegungssätze der ganzen rationalen Zahlen, bzw. der Ideal in algebraischen Zahlkörpern auf Ideale in beleibigen Integritäts-, allgemeiner Ringbereichen.*

[10] À propos de l'axiomatisation du concept d'anneau par Fraenkel, voir Corry (1996, §4.4 et 4.5).

1. associativité de l'addition : pour tous a, b et c, $(a+b)+c=a+(b+c)$;
2. commutativité de l'addition : pour tous a et b, $a+b=b+a$;
3. associativité de la multiplication : pour tous a, b et c, $(a\cdot b)\cdot c=a\cdot(b\cdot c)$;
4. commutativité de la multiplication : pour tous a et b, $a\cdot b=b\cdot a$;
5. distributivité de la multiplication sur l'addition : pour tous a, b et c, $a\cdot(b+c)=a\cdot b+a\cdot c$;
6. unicité de la soustraction : pour tous éléments a et b, il existe un unique élément x tel que $a+x=b$, c'est-à-dire $x=b-a$.

Noether se tourne ensuite vers les idéaux. Elle définit tout d'abord un *idéal* M de Σ comme une collection d'éléments de Σ telle que

1. si f est un élément de M et a est un élément de Σ, alors M contient $a\cdot f$.
2. si f et g sont des éléments de M, alors M contient la différence $f-g$. De plus, pour tout entier n, M contient nf .

Diverses notions relatives aux idéaux sont par la suite présentées, à commencer par celles d'union et d'intersection d'idéaux[11]. Soient M et N deux idéaux. L'union (M,N) de M et N est l'ensemble des éléments $a+b$ où a est dans M et b est dans N. L'*intersection* $[M,N]$ de M et N est l'ensemble des éléments $a\cdot b$ où a est dans M et b est dans N.

Par ailleurs, un idéal M a une *base finie* s'il existe un nombre fini d'éléments f_1, f_2,…, f_n tels que $M=(f_1,\ldots,f_n)$, c'est-à-dire que pour tout f dans M, $f=a_1f_1+\ldots+a_nf_n$ où $a_1,\ldots,a_n$ sont des éléments de Σ. Un anneau Σ satisfait la *condition de finitude* si tout idéal de Σ possède une base finie.

[11] Noether parle en fait du plus grand commun diviseur et du plus petit commun multiple. Plus généralement, elle n'utilise pas la terminologie et la notation ensembliste. Par exemple, elle dit que l'élément f divise l'idéal M pour affirmer que f appartient à l'idéal M et écrit $f\equiv 0\,(M)$ au lieu de $f\in M$. À ce sujet, voir Corry (1996, 230).

Noether précise alors qu'elle se restreindra aux anneaux ayant une base finie. Elle traite donc cette condition comme un axiome supplémentaire que les anneaux considérés devront satisfaire. Immédiatement après avoir fait cette remarque, elle démontre l'équivalence de la condition de base finie et de la condition de chaîne ascendante. La condition de chaîne ascendante stipule que, étant donnée une suite croissante d'idéaux $M_1 \subset M_2 \subset \ldots$, il existe un n tel que $M_n = M_{n+1} = M_{n+2} = \ldots$. Dans la terminologie contemporaine, Noether travaille donc avec des anneaux *noetheriens*[12].

Une représentation $M = [B_1, \ldots, B_k]$ est une *représentation réduite* si (i) aucun des idéaux B_i de la représentation ne contient l'intersection d'autres idéaux de la représentation et (ii) si aucun idéal B_i ne peut être remplacé par un autre idéal le contenant proprement[13].

Noether est alors en mesure de démontrer le lemme 1 (*Hilfsatz I*), lequel affirme que si un idéal peut être représenté comme l'intersection d'un nombre fini d'idéaux, alors il existe au moins une représentation réduite de cet idéal[14].

Soient M un idéal et $M = [B_1, \ldots, B_n]$ une représentation de M en tant qu'intersection des idéaux $B_1, \ldots, B_n$. Noether considère, pour un i donné, l'idéal U_i formé de l'intersection de tous les idéaux de la représentation qui sont distincts de B_i. Autrement dit, $U_i = [B_1, \ldots, B_{i-1}, B_{i+1}, \ldots, B_n]$. Trivialement, $M = [U_i, B_i]$ est une représentation de M.

Deux cas se présentent. D'une part, s'il n'existe aucun idéal contenant proprement B_i qui puisse remplacer B_i dans la représentation $M = [U_i, B_i]$, alors la représentation $M = [U_i, B_i] = [B_1, \ldots, B_n]$ est par définition réduite.

D'autre part, s'il existe un idéal $B_i^{(1)}$ qui contient proprement B_i et qui peut le remplacer dans la représentation de M, alors le même raisonnement peut être appliqué à cet idéal $B_i^{(1)}$. Ainsi, s'il n'existe

[12] Dans l'article de 1926, Noether imposera quatre axiomes supplémentaires, ceux qui définissent la notion d'anneau de Dedekind. Voir Noether (1927, 26-27).

[13] Un idéal M *contient proprement* un idéal N si M contient tous les éléments de N et s'il existe au moins un élément de M qui n'est pas dans N.

[14] La formulation de la démonstration qui suit s'inspire de celle de Corry (1996, 230-231).

aucun idéal contenant proprement $B_i^{(1)}$ qui peut le remplacer dans la chaîne, alors celle-ci est réduite. Sinon, il existe un idéal $B_i^{(2)}$ qui contient $B_i^{(1)}$ qui peut le remplacer dans la chaîne.

L'application répétée de ce raisonnement générera une représentation réduite ou une chaîne infinie d'idéaux $B_i, B_i^{(1)}, \ldots$ telle que, pour tout k, $B_i^{(k)}$ est contenu proprement dans $B_i^{(k+1)}$. Or, l'existence d'une telle chaîne infinie contredit la condition de chaîne ascendante. Le lemme est donc démontré.

Cette démonstration a comme particularité de n'utiliser que les propriétés d'inclusion de la collection des idéaux. En effet, Noether n'utilise que le langage de la théorie des anneaux – principalement la condition de chaîne ascendante – et fait totalement abstraction de la nature des éléments et des propriétés de ceux-ci. Contrairement au théorème de factorisation des polynômes par exemple, le lemme 1 n'est pas dérivé de résultats connus sur les nombres.

Cette particularité de la démonstration tient à ce que Noether travaille directement avec les axiomes. Ce faisant, elle établit que la portée de la méthode axiomatique ne se restreint pas à la dimension logique des théories mathématiques. À titre de comparaison, chez Frege, la méthode axiomatique offrait une réponse à la question de la justification des propositions arithmétiques. Quant à Hilbert, il voyait dans la méthode axiomatique un outil de clarification de théories mathématiques. Or, cette clarification est essentiellement logique dans la mesure où, en plus de la question de la justification des propositions géométriques, elle permet de traiter celles de l'indépendance et de la cohérence des axiomes[15].

En utilisant la condition de chaîne ascendante, Noether parvient dans la suite de « *Idealtheorie in Ringbereichen* » à démontrer d'autres propriétés de factorisation. Par exemple, le théorème XIII affirme que, pour toute représentation réduite d'un idéal en tant qu'intersection d'idéaux irréductibles, les idéaux irréductibles associés aux idéaux premiers isolés sont uniquement déterminés. Ce théorème est important puisqu'il constitue bien plus qu'une généralisation de résultats connus pour les nombres ou encore les polynômes. Il s'agit

[15] Pour un traitement du rôle de la méthode axiomatique dans le développement de théories mathématiques, voir Schlimm (2006, 2008, 2011).

effectivement d'un théorème totalement nouveau que l'étude des nombres ou des polynômes n'aurait pas permis de dévoiler. Pour citer Corry,

> This result is one of the important innovations of Noether's article. In fact, neither the invariance of the associated prime ideals nor that of the isolated factors had appeared in the works of Lasker. Macaulay, on the other hand, had introduced the isolated ideals but as a concept built on properties of the spread of a variety. Noether's result is therefore not only a generalized formulation of the known theorem on polynomials, but in fact a new result which could not have been attained in the particular case of polynomial rings. (1996, 234)

Comme l'illustrent les considérations ci-dessus, Noether conféra à la méthode axiomatique un rôle nouveau par rapport à Hilbert : celui d'outil mathématique permettant de démontrer abstraitement des résultats.

Conclusion

La méthode axiomatique subit des changements fondamentaux durant la période de l'histoire des mathématiques que couvre la crise des fondements.

Chez Frege, les axiomes étaient conçus comme des vérités ne requérant aucune justification et dont la vérité a pour origine l'intuition spatiale. Ils constituent pour cette raison le fondement de la vérité d'autres propositions.

Avec *Grundlagen der Geometrie*, Hilbert s'éloigna de la conception frégéenne. Hilbert utilise plutôt la méthode axiomatique comme outil de clarification des théories mathématiques dans la mesure où elle permet de dévoiler leur structure logique.

Dans leur correspondance de 1899-1900, Frege s'opposera à cette conception de la méthode axiomatique. Premièrement, selon Frege, un axiome ne peut être formulé et avoir un sens que si les concepts auxquels il réfère sont clairement définis. Hilbert considère plutôt que ce sont les relations entre concepts qu'expriment les axiomes qui définissent les concepts en question. Deuxièmement, Frege considère que la cohérence est une conséquence triviale de la vérité des axiomes. Chez Hilbert, la cohérence du système d'axiomes est plutôt le critère de leur vérité, mais aussi de l'existence des objets mathématiques qu'ils définissent.

La méthode axiomatique sera utilisée d'une toute autre façon par Emmy Noether dans le cadre de ses travaux sur le théorème de factorisation unique. Comme l'illustre la condition de chaîne ascendante, les axiomes permettent de démontrer abstraitement des résultats inédits et non triviaux.

En terminant, la conception noetherienne de la méthode axiomatique apparaît comme un précédent important dans l'histoire des mathématiques. En effet, une analyse des travaux d'Alexandre Grothendieck sur les catégories abéliennes révèlerait de profondes similarités.

Bibliographie

CORRY, L., 1996. *Modern Algebra and the Rise of Mathematical Structures.* Basel : Birkhäuser Verlag. coll. « Science Networks Historical Studies ». no 17.

FREGE, G., 1884. *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*. traduit par J.L. Austin. New York : Philosophical Library.

— 1964. *The Basic Laws of Arithmetic: Exposition of the System*. traduit par Montgomery Furth, Berkeley et Los Angeles : University of California Press.

— 1969. *Les fondements de l'arithmétique*. traduit par Claude Imbert, Paris : Éditions du Seuil, coll. « L'ordre philosophique ».

— 1980. *Philosophical and Mathematical Correspondence*, sous la dir. de Gottfried Gabriel et al. traduit par Hans Kaal. Chicago : The University of Chicago Press.

GABRIEL, G. et al., 1980. Introduction to « Frege–Hilbert » In *Philosophical and Mathemaical Correspondence*. sous la dir. de Gottfried Gabriel et al. traduit par Hans Kaal.. Chicago : The University of Chicago Press, pp. 31-32.

HILBERT, D., 1956. *Grundlagen der Geometrie*. 8e édition, Stuttgart : B. G. Teubner.

— 1971. *Les fondements de la géométrie*. traduit par Paul Rossier, Paris : Dunod.

JESHION, R., 2001. « Frege's Notions of Self-Evidence. » *Mind*. 110, pp. 937-976.

NOETHER, E., 1921. « Idealtheorie in Ringbereichen. » *Mathematische Annalen*. 83, pp. 24-66.

— 1927. « Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern. » *Mathematische Annalen*. 96, pp. 26-61.

SCHLIMM, D., 2006. « Axiomatics and Progress in the Light of 20th Century Philosophy of Science and Mathematics. » In *Foundations of the Formal Sciences IV: The History of the Concept of Formal Science*, sous la dir. de Volker Peckhaus Löwe Benedikt et Thoralf Rasch. 233-253. London : College Publications, coll. « Studies in Logic Series », no 3.

— 2008. « Bridging Theories with Axioms: Boole, Stone, and Tarski. » In *New Perspectives on Mathematical Practices: Essays in Philosophy and History of Mathematics, Brussels, Belgium, 26–28 March 2007*, sous la dir. de Bart van Kerkhove. World Scientific, pp. 222-235.

— 2011. « On the Creative Role of Axiomatics: The Discovery of Lattices by Schröder, Dedekind, Birkhoff and others. » *Synthese*. 183, pp. 47-68.

La fondation logique des probabilités chez Carnap

Karine Fradet
Université de Montréal

Résumé

Carnap a eu une participation active dans les discussions entourant la crise des fondements avant de se lancer dans le projet d'établir les fondements logiques des probabilités, ce qu'il présente dans « Logical Foundations of Probability » (1950). Il a montré que les querelles entre l'interprétation des probabilités comme état du monde et comme état de connaissance de l'observateur étaient vaines puisqu'elles ne portaient pas sur le même concept. Pour lui, la logique des probabilités est la logique inductive : une probabilité décrit une relation logique, soit le degré de confirmation d'une hypothèse selon des données observées. Deux ans plus tard, dans « The Continuum of Inductive Methods » (1952), il enrichit son système et démontre qu'il existe une quantité infinie de méthodes inductives. Nous replacerons l'interprétation de Carnap parmi celles discutées à son époque et discuterons de certaines implications.

Carnap occupe un point central dans le développement de la logique inductive. D'un côté, il a montré que les querelles entre l'interprétation des probabilités comme état du monde et comme état de connaissance de l'observateur étaient vaines puisqu'elles ne portaient pas sur le même concept. Ensuite, il s'est penché sur un de ces deux concepts, les probabilités inductives, et a jeté les bases de la logique inductive dans *Logical Foundations of Probability* (1950). Il a ensuite systématisé les différentes méthodes inductives dans *The Continuum of Inductive Methods* (1952) où il présente les différentes méthodes en compétition comme faisant partie d'un système, chaque méthode étant un point sur le continuum des méthodes inductives.

Nous retracerons ici les grandes lignes de ces développements. Nous verrons tout d'abord comment Carnap tranche la querelle entre les tenants de l'interprétation fréquentiste des probabilités et ceux soutenant une interprétation épistémique en éclaircissant les fondements de cette querelle. Ensuite, nous esquisserons la conception laplacienne des probabilités qui est un bon exemple d'interprétation inductive des probabilités. Nous nous pencherons ensuite sur le travail de systématisation des méthodes inductives accompli par Carnap.

1. Deux concepts de probabilités

Carnap souligne à plusieurs reprises (par exemple, 1945b, 1955, 1962, 1973) une confusion autour du concept de « probabilité » : certains y voient une propriété empirique alors que d'autres y voient un état de connaissance (ou plutôt d'ignorance) d'un observateur. Si cette situation crée une confusion, c'est que les deux concepts ne sont pas différenciés : « Practically everyone will say that there is only one scientific meaning [for 'probability']; but when you ask that it be stated, two different answers will come forth. » (Carnap 1955: 1) Selon Carnap, il ne s'agit pas de deux interprétations du même concept, mais bien de deux concepts différents. Si un seul concept préscientifique avait été interprété de deux façons, il s'agirait de deux interprétations. Or, dans le cas de « probabilité », il s'agit de deux concepts préscientifiques différents qui ont donné lieu à deux catégories de concepts scientifiques différents (Carnap 1945b). Il ne s'agit donc pas de deux interprétations en compétition, mais de deux concepts distincts qui ont chacun leur importance et leur place dans les sciences.

Le premier de ces concepts préscientifiques est celui qu'il nomme probabilité$_1$ ou probabilité inductive. Il s'agit d'un concept de nature logique et qui réfère au degré de confirmation d'une hypothèse *h* par rapport à un ensemble de données *e*. Selon Carnap, les premiers écrits sur les probabilités traitent de cette conception inductive des probabilités, par exemple dans le *Ars conjectandi* de Bernouilli (1713) ou la *Théorie analytique des probabilités* de Laplace (1814). Au 20^{e} siècle, on la retrouve chez des auteurs comme Keynes (1921) et Jeffreys (1939 [1998]).

Le deuxième concept préscientifique est celui qu'il nomme probabilité$_2$ ou probabilité statistique. Il s'agit d'un concept de nature empirique qui réfère à la fréquence relative d'une certaine propriété dans une population. Cette conception des probabilités a pris une place importante en sciences depuis la seconde moitié du 19^{e} siècle et a été développée par des auteurs comme Fisher (1922), von Mises (1928 [1957]) et Reichenbach (1935).

Alors que les probabilités statistiques font référence aux propriétés physiques d'un objet ou d'un système (par exemple, la composition matérielle d'un dé ou la distribution d'un trait dans la population), les probabilités inductives sont relatives à l'information disponible pour un observateur (par exemple, savoir qu'un dé a la forme d'un cube sans connaître la distribution de la masse à l'intérieur). Carnap insiste sur le fait que les deux concepts sont tout aussi objectifs. Les allégations de subjectivité contre le concept de probabilité inductive seraient dues à certaines formulations subjectivistes qui portent à confusion comme celle de « degré de croyance ». Or, selon lui, les auteurs qui utilisent de telles formulations soutiennent malgré tout des théories essentiellement objectivistes (Carnap 1945b).

Le développement des probabilités statistiques étant en plein essor, Carnap se penchera sur les probabilités inductives et il établira les fondations de la logique inductive dans *Logical Foundations of Probability* (1950). Il systématisera ensuite les méthodes inductives qu'il présentera comme n'étant que quelques méthodes possibles, quelques points sur un continuum, dans *The Continuum of Inductive Methods* (1952). Mais d'abord, nous nous tournerons vers Laplace qui a joué un rôle prédominant dans les réflexions de Carnap afin de voir en quoi consistait sa vision des probabilités.

2. Les probabilités inductives avant Carnap : Laplace

Laplace soutient une conception déterministe du monde. Dans un tel cadre métaphysique, l'idée de probabilités n'est possible que dans l'ignorance des conditions initiales de ce monde et des lois qui y opèrent. L'idée même d'une probabilité ne peut donc être due qu'à l'état de nos connaissances qui sont incomplètes : « La probabilité est relative en partie à [notre] ignorance, et en partie à nos connaissances. » (Laplace 1814: iv) Il définit la probabilité d'un événement comme suit :

> La théorie des hasards consiste à réduire tous les événements du même genre, à un certain nombre de cas également possibles, c'est-à-dire, tels que nous soyons également indécis sur leur existence ; et à déterminer le nombre des cas favorables à l'événement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles, est la mesure de cette probabilité qui n'est ainsi qu'une fraction dont le numérateur est le nombre de cas favorables, et dont le dénominateur est le nombre de tous les cas possibles. (Laplace 1814: iv)

On doit donc posséder une liste des cas également possibles, soit ceux où nous sommes également indécis ou pour lesquels « (…) rien ne porte à croire que l'un de ces cas doit arriver plutôt que les autres (…) » (Laplace 1814: 179) et diviser le nombre de cas favorables par le nombre total de cas possibles.

$$P(x) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

eq. 1) La probabilité d'un événement x chez Laplace.

Sa conception des probabilités est clairement celle d'un degré de confirmation, ce qui est sans équivoque un peu plus loin lorsqu'il donne l'exemple des trois urnes.

> Supposons, par exemple, que l'on ait trois urnes, A, B, C, dont l'une ne contienne que des boules noires, tandis que les deux autres ne renferment que des boules blanches. On doit tirer une boule de l'urne C, et l'on demande la probabilité que cette boule sera noire. Si l'on ignore quelle est celle des trois urnes, qui ne renferme que des boules noires, en sorte que l'on n'avait aucune raison de croire qu'elle est plutôt C, que B ou A ; ces trois hypothèses

> paraitront également possibles ; et comme une boule noire ne peut être extraite que dans la première, la probabilité de l'extraire est égale à un tiers. Si l'on sait que l'urne A ne contient que des boules blanches, l'indécision ne porte plus alors que sur les urnes B et C, et la probabilité que la boule extraite de l'urne C sera noire, est un demi. Enfin, cette probabilité se change en certitude, si l'on est assuré que les urnes A et B ne contiennent que des boules blanches. (Laplace 1814: v-vi)

Si le contenu en tant que tel de l'urne C n'est pas une question de probabilité (la situation est déterminée puisqu'elle contient déjà soit uniquement des boules noires, soit uniquement des boules blanches) on peut tout de même parler de la probabilité d'y piger une boule noire relativement à l'ignorance de la personne qui fera la pige. Si nous apprenons que l'urne A ne contient que des boules blanches, le nombre de cas possibles passe alors de 3 à 2. Grâce à cette information, la probabilité de piger une boule noire dans l'urne C passe de 1/3 à 1/2 puisqu'il n'y a maintenant plus que deux cas possibles qui correspondent à l'emplacement probable des boules noires. Finalement, si nous apprenons que l'urne B contient également des boules blanches, il n'y a plus qu'un seul cas possible où peuvent se trouver les boules noires et la probabilité est alors de 1. Encore une fois, ce ne sont pas les propriétés du système physique qui sont modifiées mais les données qui sont disponibles pour l'observateur.

Relativement à cet exemple, Laplace fait une affirmation qui semble aller contre le bon sens dans le calcul des probabilités mais qui sera centrale chez Carnap : « On voit par cet exemple l'influence des événements passés sur la probabilité des événements futurs. » (Laplace 1814: ix-x) Nous y reviendrons plus loin.

L'exemple de l'urne de Laplace illustre également le principe de la raison insuffisante ou principe d'indifférence : bien que la probabilité « réelle » de piger une boule noire dans l'urne C soit entièrement déterminée (elle est soit de 0, soit de 1), par manque d'information et par la symétrie de la situation, on doit attribuer une probabilité d'un tiers à l'événement qui consiste à piger une boule noire dans l'urne C.

Un exemple tiré de Carnap (1955) illustre bien le principe d'indifférence. Un dé est présenté à trois observateurs. Au premier observateur X_1, on donne l'information e_1 que le dé a la forme d'un cube régulier. Par la condition de symétrie et puisque le dé possède six faces, X_1 en conclut que la probabilité de rouler un as au prochain jet est de 1/6. À l'observateur X_2, on fournit en plus l'information e_2 comme quoi le dé est biaisé en faveur d'un des côtés. Sans autre information, X_2 doit également conclure que la probabilité de rouler un as au prochain jet est de 1/6 puisqu'il n'a aucune raison de favoriser un côté plutôt que les autres : toute présomption tant qu'au côté biaisé serait purement arbitraire. Finalement, à X_3, on ajoute l'information e_3 comme quoi le dé est biaisé en faveur de l'as. Celui-ci peut alors conclure que la probabilité de rouler un as au prochain jet est supérieure à 1/6. Dans les trois cas, les propriétés physiques de l'objet ne changent pas, ce n'est que l'information disponible pour les observateurs qui varie et la probabilité est une inférence logique. « Inductive probability characterizes a hypothesis relative to available information; this information may differ from person to person and vary for any person in the course of time. » (Carnap 1955: 4)

Le dernier principe central chez Laplace dont nous discuterons ici est la règle de succession.

> On trouve ainsi qu'un événement étant arrivé de suite, un nombre quelconque de fois ; la probabilité qu'il arrivera encore la fois suivante, est égale à ce nombre augmenté de l'unité, divisé par le même nombre augmenté de deux unités. (Laplace 1814: xiii)

Cet extrait traite d'un événement s'étant produit à chaque reprise mais plus généralement, la probabilité qu'un événement se produise la fois suivante est égale au nombre de fois qu'il s'est produit dans le passé plus 1, divisé par le nombre total de fois où il aurait pu se produire plus 2[1].

3. Les méthodes inductives

Même si on élimine les désaccords à propos de la *vraie* signification du concept de probabilité et qu'on accepte qu'on puisse en parler de façon inductive, il demeure un vaste choix de méthodes

[1] Voir l'équation 3 ci-dessous.

inductives et il n'y a pas de consensus sur l'identité de celle qui s'avère être la *bonne* méthode. Dans cette section, nous verrons ce problème du choix de la méthode inductive à l'aide d'un exemple tiré de l'article « Statistical and Inductive Probability » (Carnap 1955). Nous discuterons ensuite du travail de systématisation des différentes méthodes inductives effectué par Carnap dans *The Continuum of Inductive Methods* (1952) et verrons quatre exemples de ces méthodes telles que définies par le paramètre λ qui permet de les distinguer.

3.1. Le problème du choix de la méthode inductive

L'exemple suivant illustre le problème du choix d'une méthode inductive. Le tableau 1 ci-dessous résume une situation où quatre balles sont tirées d'une urne qui contient des balles blanches et des balles noires dans une proportion inconnue.

On peut s'intéresser à deux sortes de résultats. On peut d'abord regarder les différentes *distributions individuelles*, soit l'ensemble des permutations possibles. Pour quatre piges, cela donne seize distributions individuelles (le résultat « noire, noire, blanche, noire » en est un exemple). On peut également s'intéresser aux différentes *distributions statistiques*, soit le nombre de balles noires et de balles blanches qui feront partie du décompte final, peu importe l'ordre dans lequel elles ont été pigées. Il y a ici cinq distributions statistiques dont le résultat « trois noires, une blanche » est un exemple.

Dans le tableau 1, la colonne « méthode 1 » illustre une situation où le principe d'indifférence est d'abord appliqué aux distributions individuelles. Puisqu'elles sont au nombre de 16, la probabilité initiale de chacune est de 1/16. Pour connaître la probabilité d'une distribution statistique, on peut additionner la probabilité de chacune des distributions individuelles qui forment cette distribution statistique. Par exemple, une seule distribution individuelle correspond au critère de la distribution statistique « quatre noires », celle-ci a donc une probabilité de 1/16 ; par contre, quatre distributions individuelles répondent au critère de la distribution statistique « trois noires, une blanche », celle-ci a donc une probabilité de 4/16.

#	Distributions statistiques		Distributions individuelles	Méthode 1	Méthode 2	
	Nb noirs	Nb blancs		Prob. initiale des distr. individuelles	Prob. initiale des : distr. statistiques	distr. individuelles
1	4	0	●●●●	1/16	1/5	1/5 = 12/60
2	3	1	●●●○	1/16	1/5	1/20 = 3/60
			●●○●	1/16		1/20 = 3/60
			●○●●	1/16		1/20 = 3/60
			○●●●	1/16		1/20 = 3/60
3	2	2	●●○○	1/16	1/5	1/30 = 2/60
			●○●○	1/16		1/30 = 2/60
			●○○●	1/16		1/30 = 2/60
			○●●○	1/16		1/30 = 2/60
			○●○●	1/16		1/30 = 2/60
			○○●●	1/16		1/30 = 2/60
4	1	3	●○○○	1/16	1/5	1/20 = 3/60
			○●○○	1/16		1/20 = 3/60
			○○●○	1/16		1/20 = 3/60
			○○○●	1/16		1/20 = 3/60
5	0	4	○○○○	1/16	1/5	1/5 = 12/60

Tableau 1 : Tableau des distributions individuelles et statistiques pour quatre piges. Tiré de (Carnap 1955: 9).

La colonne « méthode 2 » représente une méthode où le principe d'indifférence est d'abord appliqué aux distributions statistiques. Puisqu'elles sont au nombre de 5, la probabilité initiale de chacune est de 1/5. On divise ensuite cette probabilité par le nombre de distributions individuelles qui peuvent former une certaine distribution statistique pour en connaître la probabilité. Par exemple, la probabilité de la distribution individuelle « noire, noire, noire, noire » est de 1/5 puisqu'elle est la seule à pouvoir former la distribution statistique « quatre noires » ; la probabilité de la distribution individuelle « noire, noire, blanche, noire » est de 1/20 puisqu'il y a quatre distributions individuelles pouvant former la distribution statistique « trois noires, une blanche ».

Notons que ces considérations ne font de sens que si nous parlons bien de probabilités inductives. Dans le cas de probabilité statistiques, s'il s'agit d'une propriété du système que la probabilité de tirer une boule noire soit, par exemple, de 50%. La méthode 1 est celle à laquelle nous sommes habitués, c'est-à-dire que la probabilité de n'importe quelle distribution individuelle est égale au produit des probabilités pour chaque événement, dans ce cas $\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}$, ce qui nous donne 1/16. La méthode 2 ne fait aucun sens si on connaît d'avance la probabilité (au sens fréquentiste) de tirer une boule noire.

Peut-on faire cette même analyse en probabilités inductives ? Après tout, avant même de tirer la première balle, puisqu'il y a deux cas possibles et que nous n'avons aucune raison d'en favoriser un, le principe d'indifférence nous dicte d'attribuer une probabilité de 1/2 aux deux événements. Ceci étant déterminé, on pourrait poursuivre notre démarche avec nos calculs habituels tels que celui décrit ci-dessus. En d'autres mots, nous assumons dès le départ par le principe d'indifférence que la fréquence relative des balles noires est de 1/2 et nous faisons les calculs habituels à partir de cette probabilité *a priori*.

Une telle méthode ne possède toutefois pas ce que Carnap juge être la caractéristique fondamentale d'une méthode inductive raisonnable, soit celle d'apprendre de l'expérience : « Inductive thinking is a way of judging hypotheses concerning unknown events. In order to be reasonable, this judging must be guided by our knowledge of observed events. » (Carnap 1955: 12) En effet, si on fixe de façon *a priori* la probabilité de tirer une balle noire pour tous les tirages à venir sans jamais pouvoir réviser ce jugement, on ne peut pas apprendre de l'expérience au fur et à mesure que les résultats empiriques sont disponibles. De même, la méthode 1 ne permet pas de corriger la probabilité en cours de route, au fur et à mesure que les balles sont tirées. Ce défaut devient clair si nous considérons une situation où nous avons tiré 100 balles de l'urne et qu'elles sont toutes noires : selon la méthode 1, on doit toujours attribuer la probabilité de 1/2 à ce que la prochaine balle tirée soit noire.

Bien que la méthode 2 ne semble pas faire de sens du point de vue des probabilités statistiques, il s'agit d'une méthode tout à fait acceptable selon Carnap au point de vue des probabilités inductives. En effet, il est possible que l'urne ne contienne que des balles noires, que des balles blanches, ou toute autre combinaison entre ces deux extrêmes. Les probabilités sont donc d'abord réparties entre les distributions statistiques avant d'être réparties entre les distributions individuelles. La répartition entre les distributions statistiques se fait selon l'équation ci-dessous qui nous donne la probabilité que le nombre de boules blanches tirées en *n* essais soit égal à B_n.

$$P(B_n) = \frac{1}{n+1}$$

eq. 2) Principe d'indifférence appliqué aux distributions statistiques.

Cela revient à attribuer la même probabilité à chacune des distributions statistiques. Il est à noter que le dénominateur est de $n + 1$ et non de n étant donné qu'il faut compter qu'il est possible qu'une distribution statistique ne contienne aucune boule blanche. Comme nous l'avons vu, la méthode 2 répartit ensuite les probabilités selon le principe d'indifférence entre toutes les distributions individuelles qui constituent une distribution statistique.

Dans cet exemple, cette méthode nous donne le même résultat que l'utilisation de la règle de succession de Laplace qui permet de déterminer les probabilités du prochain tirage en tenant compte des résultats des tirages précédents.

$$P(X_{n+1} = B \mid B_n) = \frac{B_n + 1}{n+2}$$

eq. 3) Règle de succession de Laplace.

Cette équation nous donne la probabilité que la balle (X) au prochain tirage (n+1) soit blanche (*B*), sachant que le nombre de balles blanches dans les n tirages jusqu'ici est de B_n. Si on applique cette règle pour chacun des quatre tirages consécutivement, on obtient la probabilité indiquée dans la colonne d'extrême droite pour chacune des distributions individuelles.

Par exemple, si l'on prend la distribution individuelle « blanche, blanche, noire, blanche », on a les probabilités suivantes pour chacun des tours (soit B la propriété d'être blanc et A la propriété d'être noir).

1[er] tour : $$P\left(X_{n+1}=B \mid B_n=0\right)=\frac{0+1}{0+2}=\frac{1}{2}$$

2[e] tour : $$P\left(X_{n+1}=B \mid B_n=1\right)=\frac{1+1}{1+2}=\frac{2}{3}$$

3[e] tour : $$P\left(X_{n+1}=A \mid A_n=1\right)=\frac{0+1}{2+2}=\frac{1}{4}$$

4[e] tour : $$P\left(X_{n+1}=B \mid B_n=2\right)=\frac{2+1}{3+2}=\frac{3}{5}$$

En multipliant les probabilités ainsi calculées pour chacun des tours, on obtient $\frac{1}{2}\times\frac{2}{3}\times\frac{1}{4}\times\frac{3}{5}=\frac{1}{20}$ comme probabilité pour cette distribution individuelle.

La méthode 2 sous-tend donc une réévaluation de la probabilité inductive à chacun des tirages. Elle répond ainsi à la caractéristique d'apprendre de l'expérience qui est fondamentale chez Carnap puisqu'elle réajuste les probabilités *a priori* au fur et à mesure que des résultats empiriques sont disponibles.

Les deux méthodes présentées ci-dessus ne sont que deux méthodes à l'intérieur d'un continuum de méthodes inductives possibles. Voyons maintenant comment Carnap systématise les différentes méthodes de la logique inductive.

3.2. Le continuum des méthodes inductives

Dans *The Continuum of Inductive Methods* (1952), Carnap montre que les différentes méthodes inductives forment un système et qu'elles peuvent être classées sur un continuum. Le degré de confirmation d'une hypothèse *h* relativement à un ensemble de données *e* peut être décrit par différentes méthodes d'induction ayant chacune une fonction de confirmation $c(h,e)$[2]. Cette fonction a comme paramètres deux propositions et comme résultat une valeur comprise entre 0 et 1, soit le degré de confirmation. Ces propositions font partie d'un système de langage ayant lui-même les caractéristiques décrites dans le tableau 2 ci-dessous[3].

Symbole	**Description**	**Exemple**
N	Nombre de constantes individuelles. Exemple : $a_1, a_2, \ldots, a_N$	a_1 : première balle a_2 : deuxième balle
π	Nombre de prédicats primitifs. Exemple : $P_1, P_2, \ldots, P_\pi$	P_1 : est blanc P_2 : est lisse
$k = 2^\pi$	Nombre de propriétés Q (conjonction où chaque prédicat primitif apparaît soit nié, soit non nié).	$Q_1 : P_1 \wedge P_2$ $Q_2 : P_1 \wedge \neg P_2$ $Q_3 : \neg P_1 \wedge P_2$ $Q_4 : \neg P_1 \wedge \neg P_2$
M	Propriété moléculaire (formée d'un nombre w de propriétés Q)	$P_1 \wedge (P_2 \vee \neg P_2) \rightarrow$ $Q_1 \vee Q_2 \ (w = 2)$

Tableau 2 : Propriétés d'un système de langage $\mathfrak{L}_N^\pi$.

On peut finalement fournir une description d'état (*state-description*) pour chacun des systèmes de langage qui consiste en une conjonction de chaque prédicat (ou sa négation) couplé avec chaque individu. Par exemple, pour un système contenant un prédicat et deux individus, les différentes descriptions d'état sont les suivantes :

---

2 Carnap discute également des fonctions d'estimation qui ont les mêmes propriétés et qui donnent le même résultat que la fonction de confirmation correspondante pour un même paramètre λ. Nous nous restreindrons donc ici à discuter des fonctions de confirmation.

3 Ces propriétés sont détaillées dans *Logical Foundations of Probability*. Nous n'en esquissons ici que l'essentiel qui est inclus dans *The Continuum of Inductive Methods*.

#	Description d'état
1	$Pa \wedge Pb$
2	$Pa \wedge \neg Pb$
3	$\neg Pa \wedge Pb$
4	$\neg Pa \wedge \neg Pb$

Tableau 3 : Exemple de description d'état pour $\mathfrak{L}_2^1$

Dans l'exemple de la section précédente, le tableau 1 représente un système $\mathfrak{L}_N^1$, soit un nombre quelconque d'individus possédant chacun une seule propriété (*B*, est blanc) ou sa négation.

Chaque fonction de confirmation doit prendre deux facteurs en compte. Le premier est le facteur empirique, soit le nombre d'individus pour lesquels on a observé une certaine propriété, divisé par le nombre total d'individus observés dans l'échantillon. En gardant la notation établie dans la section précédente, cela nous donne $\frac{B_n}{n}$.

Le second facteur est le facteur logique, soit celui qui correspond à la probabilité établie *a priori*. Ce facteur est $\frac{w}{k}$, soit le nombre de propriétés Q décrivant la conjonction des propriétés recherchées divisé par le nombre total de propriétés Q du système $\mathfrak{L}_N^\pi$. Dans l'exemple du tableau 2 ci-dessus, il y a 4 propriétés Q possibles ($k = 2^\pi$ donc $k = 2^2$). La propriété *M* donnée en exemple nous dit que l'on s'intéresse à une balle qui soit blanche, peu importe qu'elle soit lisse ou non. Deux propriétés Q décrivent une telle situation, soit Q_1 et Q_2. Nous avons donc $w = 2$. Dans ce cas, le facteur logique est donc égal à 2/4, soit 1/2.

Voyons maintenant la valeur de ces deux facteurs pour les données présentées dans le tableau 1. Prenons le cas où les balles tirées étaient respectivement « blanche, blanche, noire et blanche ». Quelle est la probabilité que la prochaine balle tirée soit blanche ? Le facteur empirique est de 3/4 (trois balles blanches sont sorties sur un total de quatre tirages) et le facteur logique est de 1/2 (il n'y a que deux propriétés Q possibles et seulement l'une d'elles contient le prédicat avec la valeur recherchée). Comment peut-on concilier les résultats de ces deux facteurs ? De plus, nous avons vu que nous pouvons calculer la probabilité du prochain tirage à l'aide de la règle de succession de Laplace. Or, ce calcul nous donne une probabilité de 4/6. Que devons-nous faire avec tous ces résultats ?

Selon Carnap, les facteurs empirique et logique sont les deux bornes à l'intérieur desquelles doit se situer le résultat de toute méthode de confirmation cohérente. Dans notre exemple, les résultats doivent donc se situer entre 0,5 et 0,75 inclusivement, ce qui est le cas de la règle de succession qui nous donne 0,67. Ce qui différencie les méthodes d'induction entre elles, c'est le poids qu'elles accordent à chacun des deux facteurs. En d'autres mots, les différentes méthodes inductives font une moyenne pondérée (u) entre les deux facteurs (u_1 et u_2) comme dans l'équation ci-dessous. Ce qui les différencie est le poids W qu'elles accordent à chacun des facteurs.

$$u = \frac{W_1 u_1 + W_2 u_2}{W_1 + W_2}$$

eq. 4) La probabilité inductive donnée par une méthode est le résultat d'une moyenne pondérée entre les facteurs empirique et logique.

Par convention, Carnap fixe le poids du facteur empirique au nombre total d'individus observés dans l'échantillon (n). Ce choix a l'avantage de rendre compte du fait que la fiabilité du facteur empirique augmente lorsque la taille de l'échantillon augmente, le poids qu'on désire lui donner augmente donc également. Par exemple, la valeur du facteur empirique risque plus de se rapprocher de la fréquence relative (au sens fréquentiste) des balles dans l'urne si on a tiré 500 balles que si on n'en a tiré que deux.

Le poids du facteur logique est quant à lui représenté par le symbole λ. La forme générale d'une méthode de confirmation sera donc la suivante.

$$u = \frac{n\left(\frac{B_n}{n}\right) + \lambda\left(\frac{w}{k}\right)}{n+\lambda} = \frac{B_n + \lambda\left(\frac{w}{k}\right)}{n+\lambda}$$

eq. 5) Facteurs empirique et logique pondérés par n et λ respectivement.

Les méthodes inductives peuvent donc être caractérisées à l'aide d'un seul paramètre λ qui donne le ratio pour chacun des facteurs. Celui-ci peut prendre n'importe quelle valeur parmi les nombres réels incluant 0 et ∞.

4. Quatre applications

Dans cette section, nous examinerons quatre cas de figure où λ prend chacune des deux valeurs extrêmes, 0 et ∞, ainsi qu'une valeur intermédiaire λ = 2. Finalement, nous verrons la fonction *c** développée par Carnap dont la valeur n'est pas indépendante de la valeur de *k*.

4.1. λ = 2 ou la règle de succession de Laplace modifiée

Pour pallier à certaines incohérences pouvant résulter de la règle de succession de Laplace, Carnap la modifie afin qu'elle prenne en compte le nombre de prédicats. Celle-ci correspond alors à la méthode λ = 2.

$$P\left(X_{n+1} = B \mid B_n\right) = \frac{B_n + \frac{2w}{k}}{n+2}$$

eq. 6) $c_2\left(h, e\right)$. Règle de succession de Laplace modifiée.

Le facteur logique pondéré $\lambda\left(\frac{w}{k}\right)$ peut être simplifié puisque la méthode de Laplace porte directement sur les prédicats primitifs pour lesquels $\frac{w}{k}=\frac{1}{2}$ puisque $w=\frac{k}{2}$. Cette simplification nous donne alors l'équation 3. Toutefois, telle que formulée à l'équation 6, cette méthode est plus générale et permet une application sur des propriétés moléculaires sans que cela ne cause les incohérences répertoriées dans la littérature.

4.2. λ = ∞ ou la toute-puissance du facteur logique

La méthode $\lambda=\infty$ correspond à la méthode 1 du tableau 1. En simplifiant la fonction, on retrouve le facteur logique, soit le nombre w de propriétés Q correspondant à la situation recherchée relativement au nombre k de propriétés Q du système $\mathfrak{L}_N^\pi$.

$$P\left(X_{n+1}=B \mid B_n\right)=\lim_{\lambda\to\infty}\frac{B_n+\lambda\left(\frac{w}{k}\right)}{n+\lambda}=\frac{\lambda\left(\frac{w}{k}\right)}{\lambda}=\frac{w}{k}$$

eq. 7) $c_\infty(h,e)$. Simplification de la fonction de confirmation pour $\lambda=\infty$.

Le facteur empirique est entièrement absorbé par la valeur λ = ∞. Cette méthode ne tient donc aucunement compte des informations empiriques obtenues. Comme nous l'avons vu, cette méthode est par le fait même insatisfaisante pour Carnap puisqu'elle ne respecte pas le principe selon lequel on doit apprendre de l'expérience.

4.3. λ = 0 ou la règle stricte

La méthode λ = 0 est celle qui n'accorde aucun poids au facteur logique. Elle tient compte du facteur empirique sans porter aucune attention aux probabilités qui pourraient être données *a priori* ou par le principe d'indifférence.

$$P\left(X_{n+1}=B \mid B_{n}\right)=\frac{B_{n}+\lambda\left(\frac{w}{k}\right)}{n+\lambda}=\frac{B_{n}}{n}$$

eq. 8) $c_0(h,e)$. Simplification de la fonction de confirmation pour $\lambda=0$.

Une complication apparente est celle où l'échantillon est vide (n = 0) où on a alors 0 au dénominateur. Dans ce cas, il est possible d'approcher l'équation par sa limite lorsque λ tend vers 0, ce qui nous donne $\lim_{\lambda \to 0} c_{\lambda}(h,e)=\frac{w}{k}$.

Pour de très petits échantillons (n = 1 par exemple), considérer uniquement le facteur empirique n'est probablement pas la méthode la plus raisonnable. Par exemple, si on tire une pièce à pile ou face et qu'on obtient le résultat pile au premier jet, selon une telle méthode de confirmation, le degré de confirmation de l'hypothèse « le prochain jet donnera le résultat pile » est alors de 1 ! Il en va de même pour un dé régulier qu'on aurait roulé cinq fois et dont le résultat « 6 » ne serait pas encore sorti : cette méthode de confirmation nous dit que l'hypothèse « le prochain jet donnera un '6' » a un degré de confirmation de 0. Comme le remarque Carnap, ce genre de résultat n'est pas satisfaisant puisque la probabilité d'observer une propriété qui n'est pas encore représentée dans l'échantillon « (…) may be low, but it cannot be regarded as 0 because it is, after all, a *possible case*, i.e., one *logically compatible* with the given evidence. » (Carnap 1952: 42, emphase ajoutée)

4.4. λ = k ou la fonction c* de Carnap

La fonction que Carnap privilégie est la fonction de confirmation *c** où la valeur de λ est déterminée par la valeur de *k*.

$$P\left(X_{n+1}=B\,|\,B_n\right)=\frac{B_n+\lambda\left(\frac{w}{k}\right)}{n+\lambda}=\frac{B_n+w}{n+k}$$

eq. 9) Fonction *c** de Carnap où $\lambda = k$

Cette fonction remplit les critères d'une bonne fonction de confirmation puisque le facteur logique est pris en compte, tout en accordant de plus en plus de poids au facteur empirique au fur et à mesure que l'échantillon grossit (Carnap 1945a).

De plus, le poids de la grosseur de l'échantillon est pris en compte sur une base relative par rapport au nombre de propriétés observables plutôt que de façon absolue comme c'est le cas dans la règle de succession de Laplace. On accorde le même poids au facteur empirique et au facteur logique lorsque $n = \lambda$.

Par exemple, dans la règle de succession de Laplace modifiée ($\lambda = 2$), les deux facteurs ont le même poids dès que l'échantillon compte deux individus, et ce, peu importe le nombre de propriétés du système. Par exemple, prenons un premier système qui ne contient qu'une seule propriété comme dans le tableau 1 (la couleur, blanche ou noire). Prenons un second système qui contient une dizaine de propriétés, par exemple la couleur, la texture, la présence de pois, la présence d'une craquelure, la présence d'une dépression, etc. Intuitivement, un échantillon de 5 individus nous semblera avoir plus de poids relativement au nombre de propriétés dans le premier système que dans le second. La règle de succession de Laplace modifiée ne nous permet pas de rendre compte de cette intuition alors que pour la règle *c** de Carnap, le facteur empirique aura le même poids que le facteur logique seulement à partir du moment où l'échantillon comptera le même nombre d'individus qu'il y a de propriétés *Q* possibles. Si $n < k$, le facteur logique sera prépondérant ; si $n > k$, le facteur empirique sera prépondérant. Dans notre exemple, dans le premier système, les deux facteurs auront une valeur pondérale égale après un échantillon de $n = 2$ individus alors que dans le second, leur valeur pondérale sera égale après un échantillon de $n = 2^{10} = 1024$ individus.

Conclusion

En plus d'avoir mis le doigt sur la cause des ambiguïtés entourant le concept de « probabilités », Carnap a construit un système rendant compte des différentes méthodes inductives couramment utilisées et il met en évidence d'autres méthodes inductives possibles, celles-ci formant un continuum.

Toutefois, son système repose sur la logique bivalente où une propriété ne peut prendre que la valeur de « vrai » ou « faux ». Par exemple, une balle pigée dans une urne sera « blanche » ou « non-blanche ». On peut s'en sortir si une seule valeur nous intéresse : il est possible de parler d'un dé si les résultats qui nous intéressent sont « as » et « non-as », quoi qu'ici le facteur logique accorde une probabilité de 1/2 à ces deux valeurs.

Une solution serait de créer une propriété pour chaque valeur possible, par exemple « blanche », « noire », « bleue », « rouge », etc. Toutefois, Carnap précise que les propriétés d'un système doivent être indépendantes : piger une balle rouge ne doit pas impliquer qu'elle sera non-bleue. Il est possible de déformer un peu son système si on ne regarde qu'une seule propriété primitive, mais les choses se gâtent si on regarde deux propriétés (par exemple, la couleur et la texture) ou pour les cas où des propriétés auraient un nombre de valeurs différentes, par exemple la propriété couleur (5 valeurs possibles) et la propriété texture (2 valeurs possibles).

La condition d'indépendance entre les différentes propriétés rend le système très peu flexible et incapable de rendre compte de plusieurs propriétés qui pourraient intéresser l'agent épistémique. Celui-ci pourrait être intéressé à évaluer la probabilité de propositions du type « si la prochaine balle tirée est rouge, alors elle sera lisse », ce qui exprime la dépendance entre deux propriétés.

Références

BERNOULLI, J., 1713. "Ars conjectandi", Basel.

CARNAP, R., 1945a. "On inductive logic". *Philosophy of Science* **12**(2), pp. 72-97.

— 1945b. "The Two Concepts of Probability: The Problem of Probability". *Philosophy and Phenomenological Research* **5**(4), pp. 513-532.

— 1950. *Logical foundations of probability*. Chicago, The University of Chicago Press.

— 1952. *The Continuum of Inductive Methods*. Chicago, University of Chicago Press.

— 1955. "Statistical and inductive probability". *The Galois Institute of Mathematics and Art*.

— 1962. *Logical foundations of probability*, University of Chicago Press Chicago.

— 1973. "Notes on probability and induction". *Synthese* **25**(3), pp. 269-298.

FISHER, R., 1922. "On the mathematical foundations of theoretical statistics". *Philosophical Transactions of the Royal Society of London* Series A, vol. 222, pp. 309-368.

JEFFREYS, S., 1939 [1998]. *Theory of probability*, Oxford University Press, USA.

KEYNES, J., 1921. *A treatise on probability*.

LAPLACE, P., 1814. *Théorie analytique des probabilités*. Paris : Mme. Ve. Courcier.

REICHENBACH, H., 1935. *The theory of probability*.

VON MISES, R., 1928 [1957]. *Probability, Statistics and Truth*. New York : Dover Publications.

De Kronecker à Gödel via Hilbert. Les fondements arithmétiques et une crise sans fondement

Yvon Gauthier
Université de Montréal

La crise des fondements n'a pas affecté les fondements arithmétiques du constructivisme de Kronecker, Bien plutôt, c'est le finitisme kroneckerien de la théorie de l'arithmétique générale ou polynomiale qui a permis à Hilbert de surmonter la crise des fondements ensemblistes et qui a poussé Gödel, inspiré par Hilbert, à proposer une extension du point de vue finitiste pour obtenir une preuve constructive de la consistance de l'arithmétique dans son interprétation fonctionnelle « Dialectica ».

Dans ce qu'il est convenu d'appeler la crise des fondements des mathématiques du début du XX^e siècle, la logique formelle (Frege) et la théorie des ensembles (Cantor) avaient partie liée. Kronecker et son finitisme arithmétique y échappaient par définition. Hilbert y a échappé en se réfugiant chez Kronecker. Je veux montrer dans ce qui suit que les fondements arithmétiques de la théorie des formes ou polynômes homogènes de Kronecker ont fourni les armes nécessaires à Hilbert et ses successeurs pour ne pas être assujettis aux paradoxes logiques ou ensemblistes. L'idée de fonctionnelle polynomiale que l'on trouve chez Kronecker et que l'on retrouve chez Hilbert jusqu'à Gödel permet en effet de construire l'univers arithmétique sans le secours des notions ensemblistes et sans le recours à des constructions logiques que l'on voudra éliminer, une fois qu'on les a introduites pour faciliter l'accès au domaine purement mathématique, i.e. arithmétique. Je veux donner ici une brève idée de cette problématique que j'ai amplement développée ailleurs (voir Gauthier [1] à [7]).

Kronecker et la théorie des formes d'ordre supérieur

Dans un texte bref de 1883 « Zur Theorie der Formen höherer Stufe » ([12]), Kronecker veut compléter la théorie des formes ou polynômes homogènes qu'il a élaborée dans son œuvre majeure de 1882 ([11]) « Grundzüge einer arithmetischen Theorie der algebraischen Grössen » ou « Les éléments fondamentaux d'une théorie arithmétique des grandeurs algébriques ». Dans ce texte, Kronecker a défini la construction ou la structure générale du contenu polynomial et sa décomposition (Zerlegung) en termes de la théorie de la divisibilité : la théorie comprend en effet les formes ou polynômes homogènes et les systèmes modulaires « Modulsysteme » ou systèmes de diviseurs. Je propose ici ma propre reconstruction du schème kroneckerien en utilisant le produit de convolution (produit de Cauchy) pour les polynômes

$$f \cdot g = \left(\sum_m f_m x^m\right) \left(\sum_n g_n x^n\right) = \left(\sum_m \sum_n f_m g_n x^{m+n}\right)$$

avec addition des coefficients *m* et *n*. Kronecker ([1882] : 343) dit qu'une forme *M* est contenue dans une autre forme *M'* si les coefficients de l'une sont contenus dans les coefficients de l'autre. Il formule alors des propositions générales sur l'équivalence des formes :

> Les formes homogènes linéaires qui sont équivalentes peuvent être transformées entre elles par substitution avec des coefficients entiers.
>
> (Proposition X dans Kronecker [1882] : 345)

et

> Deux formes sont absolument équivalentes lorsqu'elles se contiennent l'une l'autre.
>
> (Proposition X^0 dans Kronecker [1882] : 351)

Kronecker énonce alors son résultat principal <*Hauptresultat*> :

> Toute forme algébrique entière au sens de l'équivalence absolue de la Proposition X^0 est représentable comme produit de formes irréductibles (premières) de façon unique.
>
> (Proposition $XIII^0$ dans Kronecker [1882] : 352).

et Kronecker d'ajouter que ce résultat montre que l'association des formes algébriques entières (à coefficients entiers) à l'aide de la méthode des indéterminées conserve les déterminations conceptuelles des lois élémentaires de l'arithmétique dans le passage du domaine rationnel ou du domaine des fonctions rationnelles entières au domaine des fonctions algébriques. Mais Kronecker n'est pas entièrement satisfait et revient l'année suivante avec sa théorie des formes d'ordre supérieur et introduit le produit (Kronecker[1883] : 422)

$$\sum_{h=0}^{m} M_h U_h \cdot \sum_{i=1}^{m+1} M_{m+2} U_{m+1}$$

(où $M = M_0$, M_1 , M_2 ,…, M_{n+1} sont des quantités intégrales de domaines de rationalité successifs R et les U's sont des indéterminées); on a ainsi une forme de puissance *r contenant* le produit de formes

$$\prod_{h}^{r} \sum_{k} Mk'Vhk.$$

Kronecker souligne que cette formulation est plus générale que sa théorie de 1882. « Être contenu » ici signifie seulement que les polynômes dans les domaines de rationalité sont inclus ou contenus dans un ordre supérieur de leurs coefficients entiers. Un système modulaire va alors décomposer cette construction en polynômes irréductibles, de sorte que les notions d'inclusion et d'équivalence (inclusion réciproque) valent aussi bien pour les formes que pour les diviseurs, i.e. la décomposition en facteurs est une technique de descente tout à fait similaire à l'algorithme de division pour les entiers ou à l'algorithme euclidien pour les polynômes.

Dans cette décomposition canonique des polynômes, la descente infinie à la Fermat permet d'arriver aux polynômes irréductibles de la même manière que la preuve d'Euclide sur la divisibilité des nombres composés par les nombres premiers. La version kroneckerienne de la décomposition canonique des polynômes repose sur les formules

$$\prod_{h=1}^{r} M_k U_{hk}$$

et

$$\prod_{i=j+k} c_i = \sum_{j+k=i} a_j b_k$$

avec $j = (0, \ldots, m)$ et $k = (0, \ldots, n)$; nous pouvons lire cette dernière formule du point de vue de la divisibilité avec $a^{p-1} \equiv 1 (\mathrm{mod}\ p)$

$$\prod_{i=1}^{m+n} (1 + c_i x_i) = \sum_{i=0}^{m+n} \left(c_i x^{m+n-1}\right) = \sum_{m+n=1} (a_m b_n) \, .$$

La généralisation de Kronecker utilise le produit de convolution pour les polynômes

$$\sum_h M_h U_h \cdot \sum_i M_{m+i} U^{i-1} = \sum_k M_k' U^k$$

où les M's sont des formes entières et les U's des indéterminées comme dans le produit introduit plus haut :

$$\prod_h \sum_k M_k' U_{hk}$$

et le "contenu" du produit peut alors être exprimé par

$$\sum_k M_k' U^k = (M_k M_{m+1})^k + (M_k M_{m+1})^{k-1} + (M_k M_{m+1})^{k-2} + \ldots + (M_k M_{m+1})$$

dans l'ordre décroissant des puissances de la somme polynômiale. Cette combinaison linéaire par le produit de convolution et la descente (finie) des puissances montre simplement que les formes rationnelles entières génèrent des formes algébriques entières, c'est-à-dire des entiers algébriques. Ce que nous trouvons dans le texte de 1883, c'est la simple généralisation aux ordres (*Stufe*) supérieurs de la théorie des formes de 1882 qui comprend à la fois la théorie des systèmes modulaires et la théorie des polynômes. Le principe d'équivalence des formes énoncé en 1882 est valide en toute généralité et la notion de contenu ou d'inclusion « *Enthalten-sein* » s'y déploie dans la construction « *Bildung* » des formes entières qui en quelque sorte remplissent totalement l'univers arithmétique (Kronecker[1883] : 423). À mon sens, cette construction des formes d'ordre supérieur que j'appelle fonctionnelles polynomiales préside à l'idée de fonctionnelle chez Hilbert.

Hilbert et l'introduction de la notion de fonctionnelle.

C'est dans son texte de 1926 ([9]) « *Über das Unendliche* » que Hilbert introduit la notion de fonctionnelle dans sa tentative de démonstration de l'hypothèse du continu de Cantor :

$$2^{\aleph_0} = \aleph_1$$

ou dans sa forme générale

$$\forall\sigma\ (2^{\aleph_\sigma} = \aleph_{\sigma+1}).$$

Il s'agit pour Hilbert de construire une hiérarchie de fonctions de fonctions, de fonctions de fonctions de fonctions ..., c'est-à-dire de fonctionnelles pour atteindre la deuxième classe de nombres de Cantor qu'il désigne par N et qui est définie par la limite

$$\lim \omega = \varepsilon_0.$$

L'ensemble des entiers Z étant défini par

$$\lim n = \omega$$

il s'agit de faire correspondre toute fonction dans Z à un nombre de la deuxième classe de Cantor N. Les moyens élémentaires pour construire ces fonctions, nous dit Hilbert, sont la substitution des variables (*Einsetzung*) et la récurrence (*Rekursion*) qui consiste à dériver la valeur d'une fonction pour $n + 1$ de sa valeur pour n – ces moyens correspondent chez Kronecker à la substitution des indéterminées (considérées comme variables) et à la descente sur les puissances construites par le produit de convolution. Hilbert continue en disant que pour les fonctionnelles il faut introduire des paramètres d'ordre supérieur, ce sont les types de variables de hauteur (*Höhe*) croissante, mais il insiste sur le fait que c'est par itération finie en accord avec le point de vue finitiste (*finite Einstellung*) que l'on construit les fonctionnelles sur les entiers qui devront correspondre à un nombre transfini de la deuxième classe, la deuxième classe étant elle-même définie par une itération transfinie de la suite des $\omega < \varepsilon_0$. Cette tentative d'associer bijectivement les fonctionnelles dans Z avec les ordinaux de la deuxième classe de nombres de Cantor pour démontrer l'hypothèse du continu a échoué, mais les successeurs de Hilbert, Gentzen et Ackermann ont voulu utiliser l'induction transfinie jusqu'à ε_0 pour démontrer la consistance de l'arithmétique de Peano. Hilbert a donc voulu, selon mon hypothèse, prolonger jusque dans le transfini la construction des fonctionnelles polynomiales ou formes d'ordre supérieur de Kronecker tout en garantissant le point de vue finitiste d'inspiration kroneckerienne. Son programme d'arithmétisation de la logique va dans le même sens et se prolonge jusqu'à Gödel dans son extension du point de vue finitiste (l'interprétation *Dialectica*) qui reprendra l'idée de fonctionnelle de Hilbert remontant, comme j'ai voulu le montrer, au finitisme arithmétique de Kronecker.

Le programme d'arithmétisation de Hilbert

L'idée de Hilbert en introduisant le symbole epsilon ε (que nous distinguons du symbole ε_0) était d'assurer le passage de l'arithmétique aux éléments idéaux de la théorie des ensembles et de l'analyse, c'est-à-dire d'assurer la consistance des mathématiques infinitaires à l'aide de l'arithmétique finitaire, la théorie Z de l'arithmétique classique (récursive primitive). Hilbert a conçu la fonction de choix transfinie pour combler le fossé entre l'arithmétique finie et l'arithmétique transfinie de Cantor (voir [10]). Mais une fois que ce niveau supérieur de l'existence mathématique est atteint, il faut redescendre à la base finie : c'est la méthode de descente <*Methode der Zurückführung*> qui consiste en une construction <*Aufbau*> et une décomposition <*Abbau*> en termes arithmétiques. Le problème de la consistance de l'arithmétique est donc une question d'arithmétique finie qu'on doit sauvegarder par une procédure d'élimination du symbole ε et des formules critiques qui y sont attachées. A la question souvent posée : "Pourquoi introduire le symbole ε si c'est pour l'éliminer tout de suite après?", la réponse est simplement : "Pour construire un royaume idéal et revenir ensuite aux fondements arithmétiques pour garantir l'édifice entier des mathématiques". La logique (avec la méthode axiomatique) n'est qu'un outil dans la mesure où elle s'occupe des inférences arithmétiques élémentaires et de leur vérité par l'extension consistante de ses méthodes, mais elle permet en même temps les inférences transarithmétiques.

Le symbole ε et son élimination.

Le premier axiome pour le symbole ε est

$$A(a) \;\rightarrow\; A(\varepsilon_x\, A(x))$$

où $\varepsilon(A)$ est une fonction logique de choix transfinie [9]. Le quantificateur existentiel est défini par

$$\exists x\; Ax \;\equiv\; A(\varepsilon_x\, A(x))$$

et le quantificateur universel par

$$\forall x\; Ax \;\equiv\; A(\varepsilon_x\, \neg A(x))$$

signifiant que la quantification universelle peut être assertée si on ne peut trouver de contre-exemple après un essai fini, *i.e.* une itération finie de la fonction de choix transfinie.

Avec l'axiome aristotélicien

$$\forall x\; Ax \;\rightarrow\; A(a)$$

et le principe du tiers exclu

$$\neg\forall x\; Ax \;\rightarrow\; \exists x\, \neg A(x)$$

ces axiomes forment le cadre axiomatique pour le symbole ε et son caractère minimal devrait ouvrir le passage de l'arithmétique à l'analyse et à la théorie des ensembles en utilisant les moyens <*Hilfsmittel*> ou le détour <*Umweg*> de la logique.

L'introduction du symbole ε requiert deux théorèmes sur les formules critiques et leur élimination : le premier théorème ε élimine les formules critiques contenant un terme t

$$A(t) \to A(\varepsilon_r\, A(r))$$

par une méthode de résolutions symboliques

$$(R) = \begin{cases} A(t_1) \to A(\varepsilon_r\, A(r)) \\ \qquad\qquad \vdots \\ A(t_n) \to A(\varepsilon_r\, A(r)) \end{cases}$$

qui reproduit la décomposition des polynômes, puisque termes et expressions sont ordonnés selon leur degré et leur rang ; le degré est le nombre (fini) maximal de termes dans une suite de termes ε et le rang d'une expression ε est le nombre (fini) maximal d'expressions dans une suite d'expressions ε. Pour les polynômes on obtient une réduction à une forme disjonctive de termes sans symbole ε, *i.e.* une expression linéaire. Le deuxième théorème ε applique le même procédé aux formules existentielles et à l'axiome d'identité. C'est le schéma d'induction qui crée un problème et requiert une nouvelle formule critique

$$A(t) \to A(\varepsilon_r\, A(r) \neq t)$$

La substitution s'effectue ici au moyen de noms de nombres ou symboles numériques <*Ziffer*> pour les termes ε et la méthode devra introduire une formulation du principe d'induction à l'aide du symbole ε. Les formules

$$A(a) \to \varepsilon_x\, A(x) \neq a'$$

et

$$a \neq 0 \to \delta(a)' = a$$

pour l'existence des successeurs et de leur récursion donnent naissance à un nouveau principe d'induction qui est formulé de la façon suivante :

> « Pour tout prédicat numérique P qui s'applique à un nombre au moins, il y a un nombre correspondant à P,

mais pour son prédécesseur, s'il y en a un, P ne s'applique pas » ([10], II, 87).

Ce principe est une conséquence directe du principe de plus petit nombre avec la fonction récursive générale μ

$$A(a) \rightarrow \mu_x A(x),$$

mais la procédure générale rappelle la décomposition polynomiale en facteurs irréductibles, *i.e.* l'algorithme euclidien pour le plus grand diviseur commun et sa généralisation par descente infinie pour les polynômes de degré n ou par la condition de la chaîne pour les anneaux (noéthériens) de polynômes.

Le principe de substitution prend la forme de substitutions partielles ou globales et les substitutions effectives de termes vont consister à trouver le polynôme de résolution en réduisant les substitutions d'instances de termes à des types fondamentaux de termes, c'est-à-dire des termes qui ne font pas partie d'autres termes. Le procédé reproduit la théorie générale de l'élimination chez Kronecker et la preuve de consistance ramènera aux formules réduites "irréductibles", comme le montre, par exemple, la preuve de consistance de Ackermann pour l'arithmétique — reproduite dans la seconde édition de Hilbert et Bernays [10], II, Supplément V, 535-555. La preuve de Ackermann repose essentiellement sur le nombre de réduction des substitutions globales <*Gesamtersetzungen*> pour les numéraux et les fonctions en recourant à la machinerie des fonctions récursives : on aboutit ainsi à une "suite normale" dont l'expression

$$n_0 \cdot 2^h + n_1 \cdot 2^{h-1} + \ldots + n_{h-1} \cdot 2 + n_h$$

pour les nombres n substitués aux termes. Le nombre de réduction a la valeur 1 ou 0 selon que la substitution globale se réduit à 0 ou à $j \neq 0$. Le nombre total de substitutions globales est 2^n lorsque le nombre de termes ε (de rang 1) dans la suite des formules est n, comme c'est le cas pour le nombre de coefficients dans un binôme, par exemple. Pour les rangs supérieurs, les équations récursives primitives suffisent

$$\psi(1, n) = 2^n$$
$$\psi(m+1, n) = 2^{n \cdot \varphi(m,n)} \cdot \psi(m, n).$$

Le deuxième théorème ε a encore affaire aux formules critiques de la seconde espèce, *i.e.* la résolution symbolique des formules existentielles en éliminant le quantificateur existentiel des formules comme

$$\exists r_1 \ \ldots \ \exists r_r \ \forall n_1 \ \ldots \ \forall n_s \ A(r_1, \ldots, n_s)$$

pour obtenir une disjonction

$$A\left(t_1^{(1)}, \ldots, t_r^{(1)}, f_1\left(t_1^{(1)}, \ldots, t_r^{(1)}\right), \ldots, f_s\left(t_1^{(1)}, \ldots, t_r^{(1)}\right)\right) \ \vee \ \ldots$$
$$\vee\, A\left(t_1^{(m)}, \ldots, t_r^{(m)}, f_1\left(t_1^{(m)}, \ldots, t_r^{(m)}\right), \ldots, f_s\left(t_1^{(m)}, \ldots, t_r^{(m)}\right)\right)$$

où les termes $t_j^{(i)}$ ne contiennent pas le symbole ε et les f_i sont des symboles de fonctions à r -arguments

$$f_1(c_1, \ldots, c_r)\,, \ \ldots \ , \ f_s(c_1, \ldots, c_r)$$

Si un axiome d'égalité est ajouté, on obtient un pur calcul des prédicats qui ouvre le chemin à une preuve de consistance dans le style de Herbrand.

Le théorème de Herbrand.

La théorie de l'élimination annonce le théorème de Herbrand sur la consistance du calcul des prédicats. Voici en bref la formulation de Herbrand. Soit A une formule en forme prenexe, par exemple

$$A \ \equiv \ \exists x \ \forall y \ \exists z \ \forall t \ R(x, y, z, t)$$

avec R sans quantificateur. On introduit deux nouvelles lettres de fonctions, f unaire et g binaire avec les termes, alors A est démontrable dans le calcul des prédicats sous la forme

$$A \equiv B\big(U_1, f(U_1), W_1, g(U_1, W_1)\big) \vee \ldots \vee B\big(U_n, f(U_n), W_n, g(U_n, W_n)\big).$$

Cette disjonction, comme celle qu'on a vue plus haut, est dérivable dans un calcul propositionnel et peut servir de critère de réfutabilité dans une interprétation négative (Voir Hilbert et Bernays [10], II, 170ss.).

La négation de A est

$$\neg A \ \equiv \ \forall x \ \exists y \ \forall z \ \exists t \ \neg B\left(x, y, z, t\right)$$

ou

$$\neg A \ \equiv \ \neg B\left(x, f\left(x\right), z, g\left(x, y\right)\right)$$

et si Herbrand a vu la consistance dans la réfutabilité dans un "champ infini" ou indéfini, Kreisel a pensé l'interprétation sans contre-exemple comme une interprétation fonctionnelle sur les types supérieurs ; les fonctionnelles récursives sont de la forme

$$Bx_1 \ \ldots \ x_n \left[F_1\left(f_1, \ldots, f_n\right), \ldots, F_m\left(F_1, \ldots, F_n\right)\right]$$

avec B ouverte. Pour une formule vraie A, nous avons

$$B\left[F\left(f, g\right), f\left(F\left(f, g\right)\right), G\left(F\left(f, g\right), G\left(F, g\right)\right)\right]$$

où les F et les G sont évidemment nos nouvelles fonctionnelles récursives sur les types.

La dernière formule A est vraie s'il n'y a pas de contre-exemple de la forme

$$\neg B\left[x, f\left(x\right), z, g\left(x, y\right)\right]$$

avec f et g comme arguments des fonctionnelles récursives F et G de type supérieur ; F et G sont continues et peuvent donc être associées à des polynômes de degré arbitraire : nous pouvons définir la composition de F et G

$$F \ \cdot \ G \ = \ \left(\sum_i F_i x^i\right) \ \left(\sum_j G_j x^j\right) \ \sum_i \sum_j \left(F_i G_j x^{i+j}\right).$$

Puisque nous ne pouvons quantifier sur toutes les fonctionnelles — par diagonalisation il y a une fonctionnelle récursive qui est distincte de toutes les fonctionnelles récursives — nous devons nous restreindre aux polynômes de degré fini et utiliser la descente sur les degrés et les hauteurs de polynômes pour retrouver une version finitiste.

Remarquons que les fonctions récursives primitives peuvent se traduire aisément en fonctions polynomiales. La chose est évidente pour les fonctions constantes initiales ; la composition et la récursion sont traitées comme un produit de convolution $G \ \cdot \ H$ pour G et H de telle sorte que

$$F(x)_{\vec{n}} = G_n\left(H_1\left(a_n\right), \ \ldots \ , \ H_p\left(a_n\right)\right)$$

Avec $H \cdot G = \sum_i \sum_j \left(G_i H_j x^{i+j}\right)$.

L'opérateur μ comme l'équivalent au principe du plus petit nombre est remplacé par la descente (finie) infinie sur les puissances décroissantes d'un polynôme de degré fini

$$F(x)_{\overleftarrow{n}} = f_0 x^n + f_1 x^{n-1} + \ldots + f_{n-1} x + f_n.$$

Selon l'idée de Hilbert d'une suite terminale de prédécesseurs pour un n donné, la descente fermatienne autorise un processus de réduction fini à la façon d'un ordre linéaire décroissant de puissances pour un polynôme donné.

L'élimination des quantificateurs.

Une autre ligne d'attaque dans le programme métamathématique de Hilbert a été conduite par Tarski et a mené à la théorie des modèles. L'élimination des quantificateurs a permis à Tarski d'obtenir une solution positive au problème de la décision pour l'algèbre et la géométrie élémentaires [13], dont l'aboutissement est une forme normale disjonctive (disjonction de conjonctions de formules atomiques) proche parente du théorème de Herbrand et du théorème de Hilbert-Ackermann pour les théories ouvertes, c'est-à-dire les théories dont les axiomes non logiques sont des formules sans quantificateurs. Ici se trouve sans doute un terrain de rencontre pour la théorie des démonstrations et la théorie des modèles — cette dernière a depuis évolué de façon indépendante sous la poussée du théorème de compacité. Mais pour arriver à son résultat syntaxique, Tarski a suivi une route similaire à la théorie de l'élimination de Hilbert (ou de Hilbert-Kronecker). Le point de départ est un système de polynômes (voir [13], 31)

$$\begin{gathered}
\alpha \equiv \alpha_0 + \alpha_1 \xi + \ldots + \alpha_m \xi^m \\
\beta \equiv \beta_0 + \beta_1 \xi + \ldots + \beta_n \xi^n \\
\gamma_1 \equiv \gamma_{1,0} + \gamma_{1,1} \xi + \ldots + \gamma_{1,n_1} \xi^{n_1} \\
\vdots \\
\gamma_r \equiv \gamma_{r,0} + \gamma_{r,1} \xi + \ldots + \gamma_{r,n_r} \xi^{n_r}
\end{gathered}$$

pour lequel est définie une fonction T pour les formules Φ de la forme

$$\left(\underset{k}{E}\xi\right)\ [a=0].$$

où $\underset{k}{E}\,\xi$ signifie "il y a exactement k valeurs de ξ" telles que $T\,(\Phi)$. est une formule sans quantificateurs équivalente. La procédure d'élimination repose ici sur le théorème de Sturm sur le nombre de racines réelles d'un polynôme entre deux valeurs quelconques $f_0\,(x)$ et $f_1\,(x)$ de la variable et se réduit à l'algorithme d'Euclide pour la détermination du plus grand diviseur commun de $f_0\,(x)$ et $f_1\,(x)$ dans le décompte des variations de signe pour le polynôme en question (qui est une équation ou une inégalité). Bien que Tarski mentionne Kronecker et nonobstant les remarques de van den Dries quant à la source kroneckerienne de la théorie de l'élimination [14], Tarski ne s'inspire pas directement de la théorie des formes (polynômes homogènes) de Kronecker. L'arithmétique générale des quantités algébriques chez Kronecker est une théorie du contenu des polynômes et Tarski ne va utiliser une notion de contenu que dans sa théorie de l'implication et de la conséquence logique. Dans ce contexte, le théorème de Sturm n'apparaît que comme un cas spécial de la théorie des diviseurs de Kronecker. Si Tarski conclut ([13], 53) que la méthode de la décision équivaut à une preuve de consistance et de complétude (pour les corps réels clos, par exemple), j'ai en vue plutôt l'auto-consistance de l'arithmétique. Mais auparavant, je résume l'idée chez Gödel d'une preuve de consistance interne comme une extension du point de vue finitiste.

La construction de Gödel

Gödel [8] a fait appel aux fonctionnelles sur tous les types finis en tant qu'objets abstraits distincts des nombres naturels (concrets) et c'est à Hilbert qu'il attribue la notion même de fonctionnelle dans une rare référence au texte hilbertien « Sur l'infini ». Remarquons cependant que Gödel réfère implicitement à la construction de Hilbert dans son texte de 1931 sur la complétude et la consistance (cf. « *Über Vollständigkeit und Widerspruchsfreiheit* », *Ergebnisse eines* *mathematischen Kolloquiums* 3, (1932) :13) où il suppose une suite transfinie de systèmes formels de types supérieurs, mais il est étonnant que dans sa preuve de consistance de l'arithmétique intuitionniste, i.e. l'interprétation *Dialectica*, Gödel se limite aux types finis jusqu'à ω. L'interprétation *Dialectica* qui est une extension du point de vue finitiste aux yeux de Gödel a fait l'objet de travaux par Spector, Howard et Kreisel dans l'esprit intuitionniste de l'induction barrée et de la récursion barrée. Quoique Gödel ait été animé par des motifs intuitionnistes, sa preuve de consistance pour l'arithmétique de Heyting peut être traduite pour l'arithmétique de Peano en transvidant son contenu tout en voulant éliminer la logique, comme chez Hilbert. Je propose une autre approche du problème de la consistance pour l'arithmétique, celle de Fermat (ou de Fermat-Kronecker) avec la descente infinie de Kronecker qui remplace l'induction de Peano et avec les indéterminées de Kronecker en lieu et place des variables fonctionnelles. L'"arithmétique générale" des polynômes (ou des formes, dans la terminologie kroneckerienne) est construite sur les suites "effinies" ou infiniment processives, dans la terminologie de Brouwer cette fois. Les suites finies sont des ensembles et le produit de convolution ou produit de Cauchy est conçu comme une application (fonction) de suites sur les suites dans N, alors que le degré d'un polynôme équivaut au type d'une formule, le motif étant celui d'une interprétation des formules en tant que polynômes. Gödel stipule, dans une phrase qui rappelle Gentzen, que la notion d'accessibilité <*Erreichbarkeit*> est un concept abstrait qui requiert une espèce de réflexion sur les constructions finies. Une telle notion est la notion d'une fonctionnelle computable de type fini sur les entiers que Gödel substitue aux notions abstraites d'assertion et de preuve en mathématiques intuitionnistes. Les formules comme

$$F' = \forall x \, \exists y \; A[x, y, z]$$

et

$$G' = \forall w \, \exists v \; B[v, w, u]$$

serviront à produire une interprétation consistante de l'arithmétique de Heyting : par exemple, nous avons

$$(F \supset G)' = \forall y, w \, \exists V Z \left[A(y, Z(y, w), x) \supset B(V(y), w, u) \right]$$

et

$$(\neg F)' \;=\; \forall y\, \exists \bar{Z} \quad \neg A\left(y, \bar{Z}\,(y)\,, x\right)$$

où x, y, v, w sont des suites finies de variables de type arbitraire, u est une suite de variables numériques et Y, V, Z et $\bar{Z}$ sont des variables du second ordre — A et B sont des formules sans quantificateurs. Ces formules généralisées forment l'interprétation fonctionnelle et les types finis sont définis par les trois clauses :

1. 0 est un type fini (le type des entiers)
2. si s et t sont des types finis, alors $s \times t$ (leur produit cartésien) est un type fini
3. si s et t sont des types finis, alors $s \to t$ est aussi un type fini.

Remarque : la troisième clause signifie que nous avons une application des fonctionnelles de type s aux fonctionnelles de type t.

Cette dernière transformation soulève des questions d'interprétation et la littérature là-dessus est abondante, mais je veux souligner seulement que cette application est pour moi un produit de convolution pour polynômes. Par l'isomorphisme de Curry-Howard, on peut identifier types et formules, en particulier, un produit de types est identifié à une conjonction de formules. J'étends cet isomorphisme en identifiant l'implication à une représentation par puissances. Les formules se rendent par

$$\exists x\ Ax \;\supset\; \exists y\ By \;=\; \left(\prod \bar{a}_0 x \;\cdot\; \prod b_0 x\right)^n$$

et

$$\forall x\ Ax \;\supset\; \forall y\ By \;=\; \prod_0^n \left(\prod \bar{a}_0 x \;\cdot\; \prod b_0 x\right)^n$$

où $\bar{a}_0$ est $1 - a$ avec les coefficients a et b et les indéterminées x. Mais ici nous avons un isomorphisme entre formules et polynômes qui semble plus naturel si on songe au fait qu'une théorie des types intuitionniste ou constructive (à la Martin-Löf) fondée sur l'isomorphisme admet des objets arbitraires (ou ensembles) dans un langage typifié dont la logique vient d'ailleurs et n'est pas motivée de façon interne. On peut penser aussi bien que l'induction sur tous les types finis a un caractère imprédicatif et ne répond pas aux exigences du finistisme, même « élargi » que propose Gödel dans son interprétation *Dialectica.*

Conclusion

J’ai voulu montrer qu’il y avait un fil continu du finitisme arithmétique de Kronecker qui va de Hilbert et ses successeurs jusqu'à l’extension du point de vue finitiste qu’a proposée Gödel dans sa preuve de consistance de l’arithmétique intuitionniste avec son interprétation fonctionnelle. J’estime que le programme de la théorie concrète (ou appliquée) des preuves pratiquée entre autres par U. Kohlenbach est dans la même veine puisque cette théorie recourt essentiellement aux ressources du théorème de Herbrand et de l’interprétation fonctionnelle de Gödel pour extraire le noyau constructif ou arithmétique des preuves en analyse classique (voir [7]). Kronecker a conçu le projet d’arithmétisation de l’algèbre abstraite - après le projet d’arithmétisation de l’analyse chez Cauchy, Weierstrass et Dedekind - qui a donné le projet d’arithmétisation de la logique de Hilbert à Gödel et qui se poursuit aujourd’hui en informatique théorique. J’ose penser que faire remonter ce fil conducteur de la notion de fonctionnelle polynomiale jusqu’à Kronecker, le premier tisseur de l’arithmétique générale dans les fondements des mathématiques, ne pourra que nous rasséréner dans l’analyse critique de fondements qui ne connaissent pas de crise.

Bibliographie

GAUTHIER, Y., 1991. *De la logique interne*, Paris : Vrin.

GAUTHIER, Y., 1994. « Hilbert and the Internal Logic of Mathematics », *Synthese*, 101, pp. 1-14.

GAUTHIER, Y., 1997. *De la logique interne. Modèles et applications*, Paris/Montréal : Diderot/Modulo.

GAUTHIER, Y., 2000. « The Internal Consistency of Arithmetic with Infinite Descent », *Modern Logic,* vol. 8, nos 1/2, pp. 47-87.

GAUTHIER, Y., 2002. *Internal Logic. Foundations of Mathematics from Kronecker to Hilbert*, Dordrecht-Boston-London, Kluwer, "Synthese Library".

GAUTHIER, Y., 2010. *Logique arithmétique. L'arithmétisation de la logique*, collection « Logique de la science », Québec : Presses de l'Université Laval.

GAUTHIER, Y., 2011. « Hilbert Programme and Applied Proof Theory », *Logique et Analyse,* 213, pp. 49-68.

GÖDEL, K., 1958. « Über eine noch nicht benütze Erweiterung des finiten Standpunkte s », *Dialectica*, 12, pp. 230-287. Voir aussi GÖDEL, K., *Collected Works*, vol. II, Oxford : Oxford University Press, 1990, pp. 271 et ss.

Hilbert, D., 1926. « Über das Unendliche », *Math. Ann.* 95, pp. 161-190, trad. par André Weil sous le titre "Sur l'infini" dans *Acta Mathematica*, vol 48 (1926), pp. 91-122.

HILBERT, D. et P. BERNAYS, 1968 et 1970, *Grundlagen der Mathematik I et II*, 2 Aufl., Berlin: Springer-Verlag.

KRONECKER, L., 1968. « Grundzüge einer arithmetischen Theorie der algebraischen Grössen », *Werke*, vol. III, ed. by K. Hensel, 5 vols., Chelsea : New York, pp. 245-387.

KRONECKER, L., 1965. « Zur Theorie der Formen höherer Stufe », *Werke*, vol. II, ed. by K. Hensel, 5 vols., Chelsea, New York, pp. 419-424.

TARSKI, A., 1951. *A decision method for elementary algebra and geometry*, 2e éd. revue, Berkeley et Los Angeles: University of California Press.

VAN DEN DRIES, L., 1988. « Alfred Tarski's Elimination Theory for Real Closed Fields », *JSL*, vol. 53, pp. 7-19.

La logique ordinale de Turing

Benoit Potvin[1]
Université de Montréal

Résumé

Dans Systems of logic based on ordinals *(1939), Turing explore les possibilités de minimiser les effets du théorème d'incomplétude pour l'arithmétique par le biais d'une logique ordinale. Nous rendons ici compte de cette recherche méconnue menée par Turing sur les fondements des mathématiques en replaçant ses apports dans le contexte actuel de la théorie de la calculabilité.*

Outre son article de 1936 *On computable numbers, with an application to the Entscheidungsproblem* – sans doute le plus fameux – Turing écrivit un autre texte, *Systems of logic based on ordinals* (1939), où il est question d'explorer les possibilités d'échapper aux conséquences du théorème d'incomplétude de Gödel pour l'arithmétique par le biais d'une *logique ordinale*. Dans une démarche sans contredit authentique, Turing obtint un résultat partiel de complétude en plus d'introduire l'*oracle*, un concept théorique important de l'informatique. Nous rendons ici compte de cette recherche méconnue menée par Turing sur les fondements des mathématiques en replaçant ses apports dans le contexte actuel de la théorie de la calculabilité et soulignant, du même fait, le caractère innovateur d'un grand homme qui ne fut peut-être pas de son époque.

« C'est l'après-demain seulement qui m'appartient.
Certains naissent posthumes. »
F. Nietzsche, *L'Antéchrist (*Avant-propos). 1895.

---

[1] L'auteur tient à remercier le professeur Yvon Gauthier pour son aide sans contredit incalculable.

1. Mise en contexte

Présenté comme étant « a profound and difficult paper »[2] par Martin Davis, *Systems of logic based on ordinals* est la thèse que Turing rédigea à Princeton sous la direction d'Alonzo Church. Suite au théorème d'incomplétude de Gödel (1931), Church et Turing avaient tous deux donné une réponse négative à l'Entscheidungsproblem mais dans des formalismes différents des fonctions récursives, à savoir respectivement le lambda-calcul (λ-calcul) et les machines de Turing. En 1937, Turing montra dans un article intitulé *Computability and Lambda-definability* (1937) l'équivalence entre le formalisme de Church et ses machines. La notion de machine développée par Turing avait cependant le mérite d'apporter une définition intuitive de la calculabilité qui n'avait pas toute la lourdeur des autres formalismes. *Systems of logic based on ordinals* est donc le troisième écrit publié par Turing et traite autant de questions mathématiques que philosophiques. C'est en 1937 que Turing quitta Newman et le King's College pour rejoindre Church à Princeton. Dans un premier temps, Church avait voulu élaborer le λ-calcul dans l'espoir de fonder solidement les mathématiques. Ce formalisme fut cependant démontré inconsistant par Kleene et Rosser (1935), deux étudiants de Church, et par le fait même, il devint possible de rendre compte de l'indécidabilité d'une manière différente de celle du théorème de Gödel (1931) et des fonctions récursives (même si la technique de preuve utilisée est la même dans les deux cas). La thèse de Turing fait donc suite à l'échec du projet fondationnel de Church. Il est intéressant de noter que Turing coucha entièrement sa thèse dans le formalisme du λ-calcul. Ce détail technique explique en partie le peu d'attention que celle-ci reçut à l'époque. Si Emil Post (1944) se référa à *Systems of logic based on ordinals*, ce ne fut que pour introduire les « degrés de Turing » sans toutefois considérer le projet d'une logique ordinale. Cette idée ne fut reprise intégralement qu'ensuite – une vingtaine d'années plus tard – par Solomon Feferman (1958) et Georg Kreisel (1960), mais d'une manière différente de Turing. À

---

[2] Turing, M. Alan. 1939. *Systems of logic based on ordinals*, in [2], p.154.

l'époque, Kleene et Rosser avaient préféré ne pas adopter le formalisme du λ-calcul dans leurs travaux, choisissant plutôt celui des fonctions récursives de Herbrand et Gödel. Il va sans dire que les recherches de Turing devenaient alors, à cause de l'utilisation du λ-calcul, beaucoup moins accessibles. Robin Gandy, unancienétudiant de Turing, écrivitdansunelettre à Max Newman (le professeur de Turing à Cambridge): « Alan considered that his paper on ordinal logics had never received the attention it deserved (he wouldn't admit that it was a stinker to read) »[3]. Toutefois, il ne faudrait pas croire que l'emploi du λ-calcul par Turing dans sa thèse ait été, de quelque façon que ce soit, contraint par Church. Tout au contraire, l'utilisation de ce formalisme était essentielle au possible succès d'une logique ordinale. Suite au théorème de Gödel, le programme de Hilbert avait été durement affecté et c'est le projet finitiste du formalisme qui était au bord du gouffre. Turing écrit : « In pre-Gödel times it was thought by some that it would be possible to carry this programme to such a point that all the intuitive judgments of mathematics could be replaced by a finite number of these rules. »[4] Le λ-calcul, de par son caractère profondément abstrait, avait la possibilité de circonscrire formellement la place nécessaire à accorder à l'intuition – notamment dans l'espoir d'identifier les formules ordinales, mais nous y reviendrons plus loin dans le texte – afin de combler les trous creusés par l'incomplétude gödelienne. On lit :

> « In consequence of the impossibility of finding a formal logic which wholly eliminates the necessity of using intuition, we naturally turn to 'non-constructive' systems of logic with which not all steps in a proof are mechanical, some being intuitive. [...]. We want it to show quite clearly when a step makes use of intuition, and when it is purely formal. The strain put on the intuition should be a minimum. »[5]

---

[3] Letter from Gandy to Max Newman, n.d. (Turing Papers, Catalogue reference A.8).
[4] Turing, M. Alan. 1939. *Systems of logic based on ordinals* in [2], p. 209.
[5] *Ibid.*, p. 210.

Le projet de Turing fait donc suite au théorème de Gödel et cherche, succédant ainsi au programme de Hilbert, à reprendre et sauver l'idée du formalisme dans la quête fondationnelle des mathématiques.

2. La logique ordinale : réaction au théorème d'incomplétude

La thèse de Turing se veut une investigation sur les possibilités d'échapper à l'incomplétude des systèmes formels par le moyen d'une *logique ordinale*. Pour Turing, une *logique* est un *système formel*. Il s'agit donc d'une tentative directe de minimiser les effets du théorème d'incomplétude de Gödel dans la mesure où celui-ci concerne précisément les systèmes formels. Le premier théorème d'incomplétude (1931) énonce qu'un système formel *P*–*P* signifie que le système contient les axiomes de Peano– ne peut être en même temps consistant et négacomplet. Sans perte de généralité, il est dit que tout système formel qui permet les opérations arithmétiques de base (+, -, ^, /), soit l'addition, la soustraction, l'exponentiation et la division, est affecté par le théorème d'incomplétude. À l'époque de Turing, il y avait peu d'informations quant aux conséquences du théorème d'incomplétude, les problèmes indécidables étant rares et la hiérarchie arithmétique n'étant pas construite. Certains mathématiciens n'hésitaient pas à se demander si plusieurs problèmes auxquels ils avaient longtemps travaillé étaient effectivement solubles. Par le premier théorème d'incomplétude, Gödel stipule l'existence d'au moins une proposition arithmétique vraie de *P* qui énonce d'elle-même son indécidabilité. Plus précisément, cet énoncé à caractère autoréférentiel joue le rôle d'un point fixe pour la propriété de ne pas être démontrable dans *P*. Cette proposition indécidable qui revêtait certainement les airs d'un artifice, du moins à l'époque, semblait difficilement avoir de réelles implications pour les mathématiques. Or, ce n'est pas l'*autoréférentialité* de la proposition qui se veut être à la source de l'indécidabilité, mais bien l'application du processus de diagonalisation. L'argument diagonal consiste de manière générale à définir une relation R (a, b) – où a, b ε A – et une propriété φ des éléments de A telles que la propriété φ ne tient que dans le cas où R (a, a) ne tient pas. Turing connaissait cette méthode, que l'on retrouve certes chez Cantor et Gödel mais aussi dans le paradoxe de Russell et le paradoxe de Richard, et c'est en fonction de celle-ci qu'il établit l'indécidabilité du problème de l'arrêt dans son article de 1936, sous le paragraphe intitulé *§8. Application of the diagonal process*[6]. Un autre problème indécidable connu et formulé par Gödel était celui de la consistance des systèmes formels, tel qu'énoncé dans le deuxième théorème d'incomplétude. À cet effet, la proposition affirmant que le système *P* est consistant (Con_p) – Con_p n'a évidemment rien d'une proposition autoréférentielle – joue, de la même manière que la proposition autoréférentielle de Gödel, le rôle d'un point fixe pour la propriété de ne pas être démontrable dans *P*. Dès lors, après avoir utilisé la méthode diagonale au profit du problème d'arrêt, Turing utilise un principe découlant directement du deuxième théorème d'incomplétude pour la construction de sa logique ordinale. Dans le premier théorème d'incomplétude, Gödel suppose la consistance de *P* : si *P* est ω-consistant (ou simplement consistant selon Rosser) alors *P* est néga-incomplet. Conséquemment, le deuxième théorème d'incomplétude énonce l'indécidabilité de la proposition arithmétique Con_p à l'intérieur de *P*, c'est-à-dire avec les seuls outils formels de *P*. Dès lors, il faut nécessairement un système au moins plus fort que *P* pour démontrer Con_p. Le principe utilisé par Turing pour sa logique ordinale, qu'il nomme *systems of logic*, est celui d'une *progression récursivement consistante*.

3. La logique ordinale : ses principes

Turing débute sa thèse ainsi :

> « The well-known theorem of Gödel shows that every system of logic is in a certain sense incomplete, but at the same time it indicates means whereby from a system L of logic a more complete system *L'* may be obtained. By repeating the process we get a sequence *L*, L_1 = *L'*, L_2 =

[6] Turing, M. Alan. 1936. On Computable Numbers, With an Application to the Entscheidungsproblem, in [2], p.132.

> L_1', […] each more complete than the preceding. A logic L_ω may then be constructed in which the provable theorems are the totality of theorems provable with the help of L, L_1, L_2, […]. We may then form $L_{2\omega}$ related to L_ω in the same way that L_ω was related to L. Proceeding in this way we can associate a system of logic with any constructive ordinal. »[7]

Autrement dit, il s'agit d'obtenir une progression de systèmes formels, les uns plus complets que les précédents. Cette recherche menée par Turing est pour le moins embryonnaire et s'intéresse avant tout, d'une manière générale, à la potentialité d'une telle méthode. Turing écrit : « The subject matter, roughly speaking, is constructive systems of logic, but since the purpose is directed towards choosing a particular constructive system of logic for practical use, an attempt at this stage to put our theorems into constructive form would be putting the cart before the horse »[8]. Turing signale ici qu'il cherche avant tout, par l'idée d'une logique ordinale, à construire un outil mathématique praticable. Le principe d'une logique ordinale, dans sa quête vers la complétude, nécessite cependant de procéder au-delà des ordinaux finis afin d'échapper à l'incomplétude gödelienne. Ceci dit, ω, le premier ordinal transfini de la théorie des ensembles transfinis de Cantor, étant défini de manière imprédicative est bien sûr ni constructible de manière finie, ni calculable à l'aide d'une machine de Turing. Le principe d'une progression récursivement consistante comme fondement d'une logique ordinale s'étend donc au-delà des limites du calculable et de la prédicativité – et de là ressort le résultat partiel de complétude obtenu par Turing – à savoir que la consistance de L_ω n'étant pas garantie, il devient possible, sinon nécessaire, de poursuivre vers $L_{\omega+1}$, $L_{\omega+2}$, $L_{\omega*2}$ et ce jusqu'au premier ordinal non-récursif de Church-Kleene (ω^{CK}). Ainsi, la logique ordinale de Turing, étant fondée sur l'impossibilité même d'échapper de manière constructive et finitiste au théorème d'incomplétude de Gödel, repose d'une part sur l'hypothèse de faire appel à l'intuition

---

[7] Turing, M. Alan. 1939. *Systems of logic based on ordinals*, in [2], p. 155.
[8] *Ibid.*, p. 161.

mathématique et, d'autre part, sur la volonté de Turing à restreindre la complétude aux problèmes dits arithmétiques ou *number-theoretic problems* (*N.T.*).

4. Les théorèmes arithmétiques ou théorèmes *Numbertheoretic* (*N.T.*)

Dans son texte, Turing s'intéresse à une classe particulière de propositions arithmétiques qu'il appelle *Number-theoretic* (*N.T.*) *theorems* ou théorèmes arithmétiques et qu'il définit de la manière suivante : « […] "for each natural number x there exists a natural number y such that $\varphi(x,y)$ vanishes", where $\varphi(x,y)$ is primitive recursive. »[9]. Notons que Turing fait référence à un ensemble de fonctions bien précis qui ne correspond généralement pas dans la littérature à ce qu'on appelle les « number-theoreticfunctions » ou fonctions arithmétiques, mais à un ensemble beaucoup plus restreint, i.e. les propositions de la forme $\forall x \exists y : [\varphi(x,y) = 0]$, où $\varphi(x,y)$ est une fonction récursive. Turing écrit : « I should assert that theorems of this kind have an importance which makes it worthwhile to give them special consideration »[10]. En effet, si Turing accorde beaucoup d'importance à cette classe c'est parce qu'il présume que plusieurs problèmes de la théorie des nombres lui appartiennent, notamment l'hypothèse de Riemann et le dernier théorème de Fermat. La classe des *N.T.* théorèmes de Turing correspond aujourd'hui à un sous-ensemble de ce qu'on appelle la classe Π_2 de la hiérarchie arithmétique de Kleene. Cette hiérarchie varie en fonction de l'alternance des quantificateurs une fois la proposition mise en forme prénexe. Autrement dit, les formules du premier ordre étant toutes logiquement équivalentes à une formule de forme prénexe, il est possible d'associer chaque proposition à une classe de la hiérarchie arithmétique. Les problèmes Π_2 ont la forme générale irréductible $\forall x \exists y F(x, y)$, sans restriction précise. Or, l'hypothèse de Riemann et le dernier théorème de Fermat sont deux problèmes qui se rapportent plutôt à Π_1, c'est-à-dire aux propositions de

---

[9] *Ibid.*, p. 163.
[10] *Ibid.*, p. 168.

la forme $\forall x F(x)$. Cette erreur n'eut toutefois aucune répercussion négative dans la mesure où le théorème d'incomplétude concerne justement la classe Π_1, soit les propositions arithmétiques vraies mais indémontrables de la forme $\forall x f(x)$. À cet effet, dans le cas du problème indécidable de l'arrêt par exemple, le complément de la proposition Σ_1 pour énoncés existentiels (les classes de la hiérarchie sont incomparables) n'est pas récursivement énumérable. Un langage (ou ensemble) est dit récursivement énumérable s'il est reconnu (i.e. accepté mais pas nécessairement décidé) par une machine de Turing. Par ailleurs, un problème, pour être décidable, doit être récursif, ce qui signifie que le langage et son complément doivent être récursivement énumérables. Dans tous les cas, Turing s'intéressa aux théorèmes *N.T.* qui appartenaient effectivement à la classe Π_2 et, de ce fait, s'il obtenait la complétude pour Π_2, il l'obtenait aussi pour Π_1. À cet effet, dans son article de 1936, Turing avait établi l'indécidabilité en fonction des machines dites *circle-free* ou non circulaires. Une machine est dite non circulaire si elle imprime sur son ruban de sortie une infinité de '*0*' ou de '*1*'. Turing écrit :

> « The behaviour of the machine may be described roughly as follows: the machine is one for the calculation of the primitive recursive function $\theta(x)$ of the number-theoretic problem, except that the results of the calculation are first arranged in a form in which the figure '*0*' and '*1*' do not occur, and the machine is then modified so that, whenever it has found that the function vanishes for some value of the argument, then '*0*' is printed. The machine is circle free if and only if $\theta(x)$ vanishes for infinitely many values of the argument. »[11]

Bien qu'on fasse habituellement référence au problème de l'arrêt, il s'agissait plutôt pour Turing du problème du non-arrêt, i.e. qu'il était impossible de déterminer si une machine allait procéder infiniment. Il est intéressant de remarquer que le problème de déterminer si une machine *M* est non circulaire correspond, dans des

[11] *Ibid.*, p. 164.

termes actuels, au problème de savoir si une fonction *f* est totale (*TOT*). Une fonction est totale si elle est définie pour toutes les entrées, c'est-à-dire pour tout son domaine de définition. Autrement dit, une machine *M* appartient à *TOT* si et seulement si $(\forall y)$ [M_w s'arrête sur *y*] (i.e. que pour tout '*y*' il existe une machine qui s'arrête sur *y* quand on lui donne le mot *w* en entrée). On voit donc que le problème consistant à déterminer si une fonction est totale appartient àΠ_2. Or, tous les problèmes *N.T.* de Turing se réduisent à un problème de *TOT*. Turing écrit :

> « We cannot say that the question about de truth of any number-theoretic theorem is reducible to a question about whether a corresponding computable function vanishes infinitely; we should have rather to say that it is reducible to the problem of whether a certain machine is circle free and calculates an identically vanishing function. »[12]

Par conséquent, si Turing obtenait la complétude pour les problèmes *N.T.*, il décidait aussi le problème de déterminer si une fonction est totale. Or, si *TOT* devenait calculable, l'argument diagonal ne serait alors plus valide et le problème de déterminer si une machine s'arrête ou non sur une entrée n'en serait plus un. Par ailleurs, que ce soit le dixième problème de Hilbert, le problème de l'arrêt ou celui de la constante Oméga de Chaitin (malgré ce que Chaitin croit), toutes les preuves d'indécidabilité se construisent par diagonalisation. Ainsi, on peut se demander, et surtout si nous étions à l'époque de Turing, si la décidabilité de *TOT* entraine l'éradication totale de l'indécidabilité. À cet égard, Turing fit un coup de maitre.

5. L'oracle

C'est dans ses travaux sur les logiques ordinales que Turing pose les fondements de la calculabilité relative, c'est-à-dire la comparaison de la complexité des problèmes incalculables en introduisant le concept d'oracle. Dit autrement, certains problèmes sont plus

[12] *Idem.*

indécidables que d'autres. La réduction ou réductibilité de Turing, outil nécessaire pour établir les degrés d'indécidabilité, sera ensuite définie par Post dans un important article de 1944, *Recursively Enumerable Sets of Positive Integers and theirDecisionProblems*[13], à la suite des propos de Turing au paragraphe 4 de sa thèse *A type of problemwhichis not number-theoretic*. Après avoir établi la classe des problèmes N.T., Turing voulut obtenir un problème qui n'appartenait pas à cette classe. De cette manière, Turing fixait la limite qu'il considérait atteignable par le biais d'une logique ordinale dans la mesure où l'oracle jouait le même rôle, ni moins fort ni plus fort, que l'intuition mathématique nécessaire au projet d'une logique ordinale. Donc, au lieu d'utiliser une fois de plus l'argument diagonal, Turing introduisit un concept original, celui d'*oracle*. Il écrit :

> « Let us suppose that we are supplied with some unspecified means of solving number-theoretic problems; a kind of oracle as it were. We shall not go any further into the nature of this oracle apart from saying that it cannot be a machine. »[14]

Turing appellera *o-machine* une machine de Turing équipée d'un oracle. Intuitivement, disons que l'oracle peut répondre à toute question précise par '*oui*' ou '1' et '*non*' ou '0'. En fait, la formalisation d'une telle machine se fait exactement de la même manière qu'une machine de Turing standard (i.e. elle s'encode de manière standard), à l'exception du fait qu'elle est équipée d'une bande supplémentaire et de trois états spéciaux. Dès lors, l'oracle ne fait pas partie de la définition même de la o-machine. On peut concevoir la solution à un problème *N.T.* particulier comme un nombre réel écrit sur la bande. Toutefois, Turing montre qu'une o-machine ne peut pas décider le problème d'arrêt pour une autre o-machine tandis qu'elle peut évidemment résoudre n'importe quel problème *N.T.* particulier. On obtient donc un problème qui n'est pas *N.T.* Autrement dit, si *TOT*

---

[13] C'est dans ce même texte de 1944 que Post énonce ce qui sera fameusement appelé le problème de Post.

[14] Turing, M. Alan. 1939. *Systems of logic based on ordinals*, in [2], pp. 166-167.

et le problème de l'arrêt étaient décidables, tous les autres problèmes ne deviendraient pas nécessairement décidables. Parmi ces problèmes au-delà de Π_2, énonçons ceux qui consistent à déterminer la complétude, la décidabilité et la consistance, pour ne mentionner que ceux-là. Bien que Turing ait cerné l'existence des différents degrés d'indécidabilité, il faudra attendre la hiérarchie arithmétique telle qu'énoncée par Kleene où, par exemple, on établit l'incomparabilité entre Σ_i et Π_i et autres corolaires. Post suggéra toutefois une façon de passer des degrés de Turing à la hiérarchie arithmétique par le théorème de Post. Pour ce faire, il faut utiliser la réduction-Turing, à savoir qu'un problème A est Turing-réductible à un problème B si, étant donné une o-machine solutionnant B, on peut aussi solutionner A. Dès lors, deux problèmes ont le même degré s'ils sont Turing-réductibles ou inter-réductibles. Les programmes indécidables varient en complexité selon le nombre de *sauts* de Turing applicables. Ceci dit, cette anticipation de Turing est pour le moins remarquable et joua un rôle dominant dans sa recherche sur la logique ordinale.

6. λ-calcul, logique ordinale et intuition.

Le λ-calcul de Church représente un moyen fort efficace pour passer de l'intuition à l'abstraction (et vice-versa), puisque les fonctions dans ce formalisme n'ont de sens que dans leur construction, i.e. leurs règles (les fonctions se construisent par *abstraction* et *application*). Ainsi, le concept fondamental du λ-calcul est que tout est fonction. Les λ-termes sont des formules bien formées (*fbf*s) du formalisme et n'ont de sens que celui que nous voulons bien leur attribuer. Évidemment, et comme le précise Turing, le sens assigné doit être invariant en fonction des conversions.[15] De ce fait, Turing assigne aux nombres entiers la signification suivante dans la formulation du calcul lambda de Church:

$$1 \rightarrow \lambda fx.f(x),\ 2 \rightarrow \lambda fx.f(f(x)),\ 3 \rightarrow \lambda fx.f(f(f(x))),\ \text{etc.}$$

[15] *Ibid.*, p. 158.

Ceci dit, définissons quelques termes utilisés par Turing et définissons plus précisément le projet d'une logique ordinale. Ce que Turing nomme une *formule logique* (ou simplement logique) a pour but de décrire un système formel particulier. Pour qu'une formule L soit une logique, L doit avoir la propriété que si L(A) conv 2 alors le λ-terme A est double. En se référant au sens donné ici-haut, on peut écrire L(A) convλfx.f(f(*x*)). A est double signifie que A(*x*) est convertible à 2 pour tous les *x*. Dire que A(*x*) est convertible à 2 signifie que pour chaque *x*, il existe un entier positif *y* tel qu'une propriété calculable θ(*x*, *y*) est vraie (ou « pour chaque nombre naturel *x* il existe un nombre naturel *y* tel que θ(*x*, *y*) s'annule »). Dès lors, tous les *N.T.T* équivalent, selon leur définition, à une proposition de type *L*(*A*) conv 2 et, inversement, toutes ces propositions correspondent à un théorème *N.T.* Une formule logique donne alors un moyen de vérifier les théorèmes *N.T.* au moyen de leurs conditions de satisfaction. Ainsi, si L est une formule logique et L(A) conv 2, alors A est double etnous savons que la proposition *N.T.* correspondante est vraie. Si *L* est une logique, on ne peut toutefois pas avoir de conditions de satisfaction pour la vérité de toutes les propositions *N.T.* Par conséquent, on détermine l'extension de L comme formant l'ensemble des propositions A pour lesquelles L(A) conv 2. Ceci dit, il est maintenant nécessaire, dans l'optique d'une logique ordinale *progressive*, de situer les formules logiques à l'aide d'une notation pour les ordinaux. Les λ-termes sont des formules ordinales dans la mesure où ils sont associés (constructivement) à des nombres ordinaux. Il n'existe pas de méthode canonique à cet effet et Turing utilise un système précis qui utilise six critères pour la détermination d'une formule ordinale[16] (il n'est pas nécessaire de les énumérer ici.). En 1938, Kleene avait justement défini une notation particulière pour les ordinaux ('O') mais Turing décida, à des fins pratiques, d'utiliser une définition quelque peu différente. À titre d'exemple, la formule ordinale Dt (ayant les caractéristiques que Dt (n, n) conv 3, Dt (n+m, n) conv 2 et Dt (n, n+m) conv 1) représente l'ordinal ω. Par conséquent, une logique ordinale est une *fbf* ⋋ telle que Λ(Ω) est une

[16] *Ibid.*, p. 173.

formule logique à chaque fois que Ω est une formule ordinale. Feferman résume bien cette idée dans un vocabulaire plus actuel : $\Lambda = L_a$ tel que a ε O, c'est-à-dire que si A est une proposition Π_2 vraie, alors L_a démontre A pour un a ε O. Or, le point est qu'il n'existe pas de procédure décisionnelle calculable, ou algorithmique, permettant de déterminer si un λ-terme est une formule ordinale (la situation serait évidemment la même avec l'utilisation des fonctions récursives) ou, dit autrement dans les termes de Kleene, 'O' n'est pas calculable. Turing connaissait dès le départ cette caractéristique particulière puisqu'il était aussi bien connu qu'aucune méthode générale permettait de déterminer si une formule du lambda-calcul avait ou non une forme normale[17]. Turing renvoie à Church ici : « there is (demonstrably) no process whereby it can be said of a formula whether it has a normal form. »[18] La détermination des formules ordinales est liée à celles de forme normale dans la mesure où si E représente une classe de formules ordinales et A n'importe quelle formule ordinale de cette classe, alors il n'existe aucune méthode par laquelle on peut déterminer si une *fbf* de forme normale appartient à E. Si Church met demonstrably entre parenthèses, c'est qu'il se rapporte à une autre procédure décisionnelle qui n'est toutefois pas mécanique, à savoir l'intuition. En effet, Church et Turing accordaient une grande place à l'intuition comme le démontre par exemple leur célèbre hypothèse – la thèse Church-Turing – sur le calculable (ce n'est bien-sûr pas un théorème). C'est dans cet esprit que le λ-calcul – l'opérateur λ est celui d'abstraction – est le formalisme le plus approprié aux fins d'une logique ordinale, dans la mesure où, au même titre que l'oracle, l'intuition permettrait de déterminer si une *fbf* est une formule ordinale et, par conséquent, d'obtenir la complétude pour les théorèmes *N.T.* Turing écrit :

> « Gödel's theorem shows that such system cannot be wholly mechanical; but with a complete ordinal logic we should be able to confine the non-mechanical steps

---

[17] Pour la forme normale, voir [2], pp. 158-159.

[18] Turing, M. Alan. 1939. *Systems of logic based on ordinals*, in [2], p.159.

> entirely to verification that particular formulae are ordinal formula.»[19]

7. Logique ordinale, complétude et invariance.

Turing obtint un résultat partiel de complétude pour les propositions Π_1 qui peut se résumer ainsi :

> Pour toute progression récursivement consistante et toute proposition Π_1 vraie $\varnothing$, il existe un a □ O où $|a| = \omega + 1$ tel que $\varnothing$ est démontrable dans L_a.

Ce résultat est partiel en ce sens que Turing espérait obtenir la complétude pour les problèmes *N.T.* (Π_2). Il écrit : « I cannot at present give a proof of this (i.e. for *N.T.* problems), but I can give a proof that it is complete as regards a simpler type of theorem than the number-theoretic theorems, viz. those of the form *θ(x)* vanishes identically where *θ(x)* is primitive recursive. »[20] À cet égard, Turing persiste à croire qu'il est possible d'obtenir la complétude pour les théorèmes *N.T*, même s'il n'est pas en mesure de le démontrer à ce moment. Feferman montra plus tard que cela n'était pas possible. Le résultat est aussi dit partiel en fonction de l'inutilité de celui-ci, dans la mesure où le problème qui consiste à reconnaitre si une formule est une formule ordinale est, tel que démontré par Kleene en 1955, déjà plus compliqué que n'importe quel autre problème *N.T.* Comme le dit Feferman :

> « Thus the demand on "intuition" in recognizing "which formulae are ordinal formulae" is somewhat greater than Turing suggests. But even in his own terms, there is a failure to test his analysis of purpose against reality. Is it a "spontaneous judgement" without any "conscious train of

[19] *Ibid.*, p. 194.
[20] *Ibid.*, p. 203.

reasoning" that leads one to recognize a complicated though computable ordering as being a well-ordering? »[21].

Turing fut d'ailleurs grandement déçu par le résultat obtenu. Précisons un point important relevé par Franzén[22] : dans le résultat de complétude de Turing, les axiomes de consistance, ajoutés les uns aux autres pour former des progressions récursivement consistantes, ne sont en aucun cas utilisés. En fait, la seule raison qui explique que $\emptyset$ soit démontrable, une fois L_ω atteint, est qu'on obtient au stade ω une définition non-standard des axiomes de L_a. Sans contredit, cela repose sur des définitions imprédicatives et est complètement inutile à des fins de calculabilité. Un autre résultat significatif est l'impossibilité d'obtenir une logique ordinale à la fois complète et invariante. Turing définit la complétude ainsi : « We say that an ordinal logic Λ is complete if corresponding to each dual formula A there is an ordinal formula Ω_A such that $\Lambda(\Omega_A, A)$ conv 2. »[23] Il s'agit donc de la complétude pour les théorèmes *N.T.* On pourrait aussi dire que la classe des formules logiques Λ(Ω) est complète si Ω parcourt l'ensemble des formules ordinales puisqu'on traiterait alors aussi toutes les propositions *N.T.* vraies. En plus de la complétude, il est souhaitable pour une logique ordinale d'obtenir l'invariance : « An ordinal logic is said to be invariant up to an ordinal *a* if, whenever Ω, Ω' are ordinal formulae representing the same ordinal less than *a*, the extent of Λ(Ω) is identical with the extent of Λ(Ω'). An ordinal logic is *invariant* if it is invariant up to each ordinal represented by an ordinal formula.»[24] La logique Λ_p (*P* pour Peano), par laquelle Turing obtint son résultat partiel de complétude, consiste à joindre successivement des propositions arithmétiques afin de *surmonter* ainsi, toujours un peu plus à chaque niveau, les effets du théorème

---

[21] Feferman, Solomon. 1988. Turing in the Land of O(z), in The Universal Turing Machine : A Half-Century Survey, p. 130.
[22] Franzén, Torkel. 2004. Inexhaustibility : a non-exhaustive Treatment , p. 190.
[23] Turing, M. Alan. 1939. *Systems of logic based on ordinals*, in [2], p. 192.
[24] *Ibid.*, p.194.

d'incomplétude. Celle-ci, étant complète pour Π_1, n'est toutefois pas invariante. Une autre logique, Λ_H (*H* pour Hilbert), est elle invariante et, par conséquent, aussi incomplète. C'est certes un résultat décevant pour Turing puisqu'il mine considérablement la portée *pratique* des logiques ordinales.

8. Conclusion

Le texte sur la logique ordinale de Turing est un maillon important dans l'histoire de l'informatique théorique du fait qu'il a contribué grandement à son développement. De plus, Turing initia un courant en fondements des mathématiques qui trouva écho chez Feferman et Kreisel. À cet égard, Feferman réussit à obtenir un résultat de complétude pour les propositions Π_2 en utilisant un principe de réflexivité appliqué à des extensions consistantes autonomes (une extension qui n'est pas une progression comme chez Turing). De plus, l'introduction du concept d'oracle est sans doute la notion qui a le plus grand impact du fait qu'elle est à l'origine de nombreuses avancées en calculabilité, suivant les recherches de Post, Sacks, Gandy et plusieurs autres. Il est intéressant de s'interroger sur la signification théorique du concept d'*oracle* dans le texte. En effet, Post soutient que le concept d'oracle est introduit en tant que *side-issue* ou idée marginale dans la mesure où ce concept n'est traité que dans le chapitre 4 :

> « In his paper on *ordinal logics*, Turing presents as a side issue a formulation which can immediately be restated as the general formulation of the recursive reducibility of one problem to another, and proves a result which immediately generalizes to the result that for any recursively given unsolvable problem there is another of higher degree of unsolvability.»[25]

Toutefois, le concept d'oracle correspondant dans le contexte du

[25] Post, Emil. 1944. Recursively Enumarable Sets of Positive Integers and their Decision Problems, in [2], pp. 310-311.

formalisme de Church à l'intuition du mathématicien, il s'agit sans contredit non pas d'un simple ajout ou d'une question secondaire de la part de Turing, mais plutôt d'un outil théorique original qui permet de juger de l'implication possible de l'intuition en théorie de la calculabilité et dans les mathématiques en général. Un autre apport significatif de ce texte est l'anticipation de la hiérarchie arithmétique et de la calculabilité relative et ce en ce qui concerne non seulement la calculabilité mais aussi la complexité de fonctions calculables. Turing écrit :

> « We might also expect to obtain an interesting classification of number-theoretic theorems according to "depth". A theorem which required an ordinal α to prove would be deeper than one which could be proved by the use of an ordinal β less than α. »[26].

Bref, pour toutes ces raisons, *Systems of logic based on ordinals* est un texte séminal de la littérature qu'il est important de considérer d'un point de vue historique, mais aussi philosophique et mathématique.

---

[26] Turing, M. Alan. 1939. *Systems of logic based on ordinals*, in [2], p. 194.

Bibliographie

COPELAND, B. J., 2004. *The Essential Turing : The ideas that gave birth to the computer age*, Oxford : Clarendon Press, pp. 125-145.

DAVIS, M., 1953. *The Undecidable : Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*, Raven Press Books Ltd.

FEFERMAN, S., 1958. *On the Strength of Ordinal Logics.* [Abstract], Journal of Symbolic Logic, March, 23(1), pp. 105-106.

FEFERMAN, S., 1958. *Ordinal Logics Re-Examined.* [Abstract] *Journal of Symbolic Logic*, March, 23(1), p. 105.

FEFERMAN, S., 1988. *Turing in the Land of O(z), in The Universal Turing Machine : A Half-Century Survey*, Oxford University Press, pp. 113-147.

FRANZÉN, T., 2004. *Inexhaustibility : a non-exhaustive Treatment*, Association for Symbolic Logic, A K Peters, pp. 153-218.

GANDY, R., 1988. *The Confluence of Ideas in 1936, in The Universal Turing Machine : A Half-Century Survey*, Oxford University Press, pp. 51-111.

GÖDEL, K., 1931. *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme*, I, Monatshefte für Mathematik und Physik, 38: pp. 173-198.

KLEENE, S. C. & ROSSER, J. B., 1935. *The inconsistency of certain formal logics, Annals of Mathematics*, 36 (3): pp. 630-636.

KREISEL, G., 1960. *Ordinal logics and the characterization of informal concepts of proof, Proceedings of the International Congress of Mathematicians* (1958), Cambridge University Press, pp. 289-299.

POST, E., 1944. *Recursively Enumarable Sets of Positive Integers and their Decision Problems*, in [2], pp. 305-337.

SYROPOULOS, A., 2008 *Hypercomputation : Computing beyond the Church-Turing Barrier*, Springer, pp. 1-24, 45-68.

TURING, M. A., 1936. *On Computable Numbers, With an Application to the Entscheidungsproblem*, in [2], pp. 116-151.

TURING, M. A., 1939. *Systems of logic based on ordinals*, in [2], pp. 155-222.

La portée épistémique de l'axiomatisation de la physique chez Hilbert

Clayton Peterson
Université de Montréal

L'objectif du présent texte est de discuter de la portée épistémique de la méthode axiomatique. Tout d'abord, il sera question du contexte à partir duquel la méthode axiomatique a émergé, ce qui sera suivi d'une discussion des motivations du programme de Hilbert et de ses objectifs. Ensuite, nous exposerons la méthode axiomatique dans un cadre plus moderne afin de mettre en lumière son utilité et sa portée théorique. Finalement, il s'agira d'explorer l'influence de la méthode axiomatique en physique, surtout en ce qui a trait à l'application de la méthode par Hilbert. Nous discuterons de ses objectifs et de l'épistémologie qui accompagnait sa vision du 6e problème, ce qui nous amènera à discuter des limites épistémiques de la méthode axiomatique et de l'entreprise scientifique en général.

1. Introduction

Les travaux des penseurs du début du 20e siècle ont grandement influencé les mathématiques, la logique et la physique moderne. Bien que plusieurs aient participé au développement des connaissances de l'époque, le programme de Hilbert a eu un impact majeur sur la conception moderne de la connaissance scientifique. La méthode axiomatique, outil central au programme de Hilbert, s'est répercutée dans différents domaines et joue maintenant un rôle épistémologique et théorique important. Ayant cela en tête, l'objectif du présent texte est d'explorer l'impact théorique et épistémologique que la méthode axiomatique a pu avoir dans le domaine de la physique.

Dans un premier temps, il s'agira d'exposer le contexte dans lequel le programme de Hilbert prend place et à partir duquel la méthode axiomatique a émergé, ce qui nous amènera à discuter du rôle de la méthode par rapport aux fondements des mathématiques. Cela fait, nous discuterons des intérêts théoriques de la méthode axiomatique à la lumière de la conception moderne de la logique. Finalement, il sera question de la répercussion de cette méthode en physique, plus précisément en ce qui a trait à l'axiomatisation de la physique chez Hilbert. Notre objectif ne sera pas d'analyser en détails le formalisme de la physique mais plutôt d'analyser les raisons qui ont pu motiver Hilbert à vouloir appliquer la méthode axiomatique à la physique. L'objectif principal sera de montrer l'intérêt et les limites épistémiques de l'application de la méthode axiomatique en physique.

2. Le programme de Hilbert

D'entrée de jeu, le programme de Hilbert est apparu en réaction à certains problèmes relatifs aux fondements des mathématiques. La connaissance mathématique, paradigme de la connaissance en vertu de sa certitude, était remise en doute à l'époque par les paradoxes de la théorie (naïve) des ensembles. Un exemple de ces paradoxes serait celui avancé par Russell, lequel s'adresse à la loi V de Frege mais plus précisément au principe naïf de la théorie des ensembles, lequel stipule que tout concept possède une extension.[1] Considérant que les paradoxes remettaient en doute la certitude de la connaissance mathématique, le programme de Hilbert s'est développé en vue de lui donner des fondements solides.[2] Croyant fermement que chaque problème mathématique possède une solution, que ce soit par une preuve directe[3] ou encore par une preuve de l'impossibilité d'obtenir une solution à partir de certaines hypothèses, l'idée de Hilbert était de

[1] Bertrand Russell, « Letter to Frege, » dans *From Frege to Gödel: A Source Book dans Mathematical Logic, 1879-1931*, éd. Jean van Heijenoort (Harvard, 1999), p. 125.

[2] David Hilbert, « The New Grounding of Mathematics, » dans *From Kant to Hilbert: A Source Book dans the Foundations of Mathematics*, éd. William Ewald (Oxford: Clarendon Press, 1922), p. 1119.

[3] ———, « From Mathematical Problems, » dans *From Kant to Hilbert: A Source Book dans the Foundations of Mathematics*, éd. William Ewald (Oxford: Clarendon Press, 1900), p. 1096.

fournir une méthode qui soit en mesure de répondre à ce présupposé.[4]

Par ailleurs, un point important à souligner est que le concept de *consistance* remplace celui d'*existence* chez Hilbert. En effet, plutôt que de définir les concepts mathématiques en fonction d'un objet mathématique qui existe (à l'instar de Frege), Hilbert est d'avis qu'un concept mathématique *existe* si et seulement si ce dernier est consistant.[5] Le concept d'existence tel qu'entendu chez Hilbert ne doit cependant pas être compris de manière métaphysique. Il suffit qu'un ensemble de concepts mathématiques ne soit pas inconsistant afin que les diverses interprétations de ce dernier puissent faire l'objet de la recherche mathématique.

Cet intérêt porté vers la consistance des concepts mathématiques se traduit chez Hilbert par un certain formalisme. Plutôt que d'insister sur l'existence des objets abstraits, Hilbert considère le *signe* comme étant l'objet mathématique.[6] Autrement dit, puisque les mathématiques se réduisent à des énoncés symboliques qui dépendent logiquement les uns des autres, il s'ensuit que la recherche mathématique se résume à l'étude des relations qui se trouvent entre ces différents symboles et énoncés.[7] En ce sens, les objets sur lesquels porte la recherche mathématique sont les symboles et les énoncés mathématiques. Cela dit, les symboles n'ont pas d'interprétation fixe. Par exemple, un symbole comme $R_1(x, y)$ peut référer à n'importe quelle relation binaire (par exemple, $x = y$, $x > y$, $x < y$, etc.). Puisque les concepts se définissent en fonction des relations qu'ils entretiennent avec d'autres concepts[8] et que les axiomes expriment ces relations, les axiomes permettent de *fixer* la référence des symboles[9]. Cependant, les axiomes ne fixent pas la référence de façon

[4] *Ibid.*, p. 1102
[5] *Ibid.*, p. 1105
[6] Hilbert (1922), *Op. cit.* note 2, p. 1121
[7] David Hilbert, « Foundations of Mathematics, » dans *From Kant to Hilbert: A Source Book dans the Foundations of Mathematics*, éd. William Ewald (Oxford: Clarendon Press, 1923), p. 1137.
[8] Gottlob Frege, *Philosophical and Mathematical Correspondence*, éd. Gottfried Gabriel (Oxford: Basil Blackwell, 1980), p. 51.
[9] *Ibid.*, p. 40

univoque. Considérant que les axiomes expriment les propriétés des relations entre les concepts, comme la symétrie, la transitivité, la densité, etc., un axiome ou un ensemble d'axiomes détermine la classe des relations auxquelles un symbole pourra référer. Par exemple, l'axiome

$$\text{(A0)} \qquad (\forall x)(\forall y)(R_1(x, y) \supset R_1(y, x))$$

ajouté à la logique du premier ordre restreint l'interprétation de R_1 aux relations symétriques, sans pour autant déterminer de quelle relation symétrique il s'agit. À partir du moment où un système formel est consistant, les interprétations de ce dernier peuvent faire l'objet de la recherche mathématique.

Selon Hilbert, les paradoxes qui minaient les fondements de la connaissance mathématique étaient principalement dus à l'application injustifiée de la logique classique à des domaines infinis. L'objection de Brower et Weyl contre l'utilisation du tiers exclu en est un bon exemple.

$$\neg(\exists x)\neg Ax \supset (\forall x)Ax$$

Alors que la logique classique peut être utilisée de façon sécuritaire à l'intérieur d'un domaine fini, il ne faut pas assumer d'emblée que ses lois s'appliquent lorsque le domaine du discours est infini. En effet, la quantification sur un domaine infini (dénombrable ou non) pose problème. Dans le fini, il est possible d'avoir une procédure décidable, c'est-à-dire une liste d'instructions claires et précises qui permettent de d'obtenir un résultat en un nombre fini d'étapes, qui permet de vérifier que nos propositions ne sont pas contradictoires. Toutefois, rien ne nous garantit que la logique classique et le tiers exclu s'appliquent lorsque le domaine est infini puisqu'il est impossible de vérifier que les propositions fonctionnent pour tous les éléments du domaine. Ce n’est pas parce que tous les objets observés jusqu’à présent ne possèdent pas la propriété A que nécessairement tous les objets du domaine ne la possèdent pas.

Afin de justifier le passage du fini à l’infini et d'éviter l'utilisation injustifiée des lois de la logique classique, Hilbert introduit une théorie de la preuve (ce que l'on nommera le *finitisme* de Hilbert). Brièvement, l’idée était de déterminer les règles qui puissent garantir l'utilisation correcte des systèmes formels. La théorie de la preuve consiste à montrer qu'une conclusion s'obtient en un nombre fini d'étapes et à partir d'un nombre fini d'hypothèses à l'intérieur d'une théorie dont le système formel contient un nombre fini d'axiomes et de règles d'inférences.[10] Autrement dit, il s'agit de montrer par le biais d'une dérivation finie, laquelle repose fondamentalement sur la logique[11], qu'une conclusion dépend logiquement d'un ensemble fini d'hypothèses et d'un ensemble fini d'axiomes.[12] La preuve est caractérisée par le fait que chaque ligne de la dérivation est soit une hypothèse, un axiome du système formel ou une proposition obtenue par règle d'inférence à partir d'une proposition précédente qui répond elle aussi à l'une de ces trois conditions.[13] De fait, il devient possible de justifier en un nombre fini d'étapes qu'une conclusion dépend logiquement d'un nombre fini d'hypothèses à l'intérieur d'un système formel (fini) donné.

Cela dit, la théorie de la preuve s'insère dans le cadre de la méthode axiomatique. Cette méthode, qui se veut principalement l'étude des relations qui se trouvent entre les énoncés et les concepts d'une théorie, possède deux objectifs principaux, à savoir montrer la consistance interne d'une théorie et montrer que le système formel permet de prouver des choses par rapport à une interprétation donnée.[14] Les axiomes d'un système sont des propositions qui expriment des relations entre des concepts ou des énoncés. La méthode axiomatique consiste à formaliser de façon syntaxique les relations qui se trouvent entre les différents symboles et concepts d'une théorie. Un système formel comprend une syntaxe (variables propositionnelles, opérateurs, symboles, constantes), un ensemble d'énoncés bien formés, un ensemble d'axiomes, un ensemble de

[10] Hilbert (1900), *Op. cit.* note 3, p. 1099

[11] *Id.* et David Hilbert, « On the Concept of Number, » dans *From Kant to Hilbert: A Source Book dans the Foundations of Mathematics*, éd. William Ewald (Oxford: Clarendon Press, 1900), p. 1138.

[12] David Hilbert, « The Foundations of Mathematics, » dans *From Frege to Gödel: A Source Book dans Mathematical Logic, 1879-1931*, éd. Jean van Heijenoort (Harvard, 1928), p. 464

[13] *Ibid.*, p. 465.

[14] Hilbert, « On the Concept of Number. » dans *From Kant to Hilbert: A Source Book dans the Foundations of Mathematics*, éd. William Ewald (Oxford: Clarendon Press, 1900), pp. 1092-1093.

règles d'inférence et un ensemble de théorèmes. Les théorèmes d'une théorie étant des énoncés qui dérivent logiquement des axiomes, il s'ensuit que les axiomes sont les propositions à partir desquelles les théorèmes d'une théorie peuvent être prouvés.[15]

D'une part, la méthode axiomatique avait pour but de montrer la consistance interne d'une théorie, c'est-à-dire montrer la consistance syntaxique du système formel. Alors que l'axiomatisation fournit une procédure de décidabilité pour les théorèmes, à savoir qu'en vertu du caractère fini du système formel et de la preuve, il est possible d'avoir une liste d'instructions claires et précises qui permettent d'énumérer tous les théorèmes en fonction de leur longueur, la preuve de consistance syntaxique requiert une procédure de décidabilité pour les non théorèmes. En effet, la preuve de consistance syntaxique d'un système K nécessite que l'on montre que celui-ci ne permet pas de dériver de contradiction:

$$\nvdash_K \bot$$

Autrement dit, il faut montrer de façon purement syntaxique que les axiomes et les règles d'inférence ne permettent pas de dériver de contradiction. Néanmoins, ces preuves de consistance peuvent être relatives à la consistance d'un autre système. En effet, dans certains cas il est possible de créer un morphisme (*mapping*) d'une théorie K à une théorie Γ de façon à montrer que si K est inconsistante, alors Γ le sera aussi (ou à l'inverse, si Γ est consistante, alors K l'est aussi). Chez Hilbert, plusieurs preuves de consistance étaient relatives à celle de l'arithmétique, comme dans le cas de l'interprétation arithmétique des axiomes de la géométrie. En d'autres termes, il s'agit d'opérer une transformation d'un système à l'autre de manière à prouver la consistance des axiomes de la géométrie en fonction de celle (présupposée) de l'arithmétique.[16] Considérant que plusieurs preuves de consistance étaient relatives à la consistance de l'arithmétique, on comprend l'importance accordée par Hilbert à la preuve de consistance de l'arithmétique, qui prend la 2e place dans la liste des 23 problèmes de Hilbert.

[15] Hilbert (1922), *Op. cit.* note 2, p.1125

[16] Hilbert (1928), *Op. cit.* note 12, p.471

Outre la preuve de consistance syntaxique du système, Hilbert cherchait aussi à montrer que celui-ci est *complet*. Cette notion de *complétude*, à savoir montrer qu'un système formel est adéquat afin de prouver des choses par rapport à une interprétation donnée (par exemple, prouver toutes les propositions vraies de la géométrie), est ce que l'on nomme la complétude sémantique d'un système. La complétude sémantique consiste à dire que toutes les propositions vraies d'une interprétation sont des théorèmes du système formel. Même si à l'époque Hilbert ne possédait pas tous les outils de la logique moderne, la complétude sémantique équivaut à montrer que si une proposition A est vraie pour une interprétation M, alors A est un théorème du système formel K:

$$\vDash_M A \Rightarrow \ \vdash_K A$$

Par surcroît, l'idée était aussi de montrer que le système formel, en plus d'être consistant syntaxiquement, était aussi consistant relativement à l'interprétation que l'on en fait. En effet, un système formel n'est pas adéquat pour représenter une théorie si l'interprétation, malgré la consistance syntaxique du système formel, est incohérente (contradictoire dans le modèle). L'interprétation d'un système axiomatique consiste à donner une signification aux symboles et aux relations, c'est-à-dire qu'il s'agit de spécifier un domaine à l'intérieur duquel les relations exprimées par les axiomes sont vraies. Ce type de considération, à savoir montrer que si une proposition A est un théorème de K, alors A sera aussi vraie pour ne interprétation M, équivaut en termes modernes à la preuve d'adéquation du système (*soundness*) :

$$\vdash_K A \Rightarrow \ \vDash_M A$$

En somme, même si les termes et stratégies que l'on retrouve en logique moderne ne se retrouvaient pas chez Hilbert, il n'en demeure pas moins que la base était déjà présente. En formalisant les relations entre les concepts d'une théorie, il devient possible d'étudier sa consistance interne et de s'assurer que le système (syntaxe) représente adéquatement l'interprétation (sémantique), et vice versa.

3. Axiomatisation et logique moderne

Malgré le fait qu'il ne soit pas toujours possible de faire des preuves purement syntaxiques de consistance[17], il n'en demeure pas moins que la notion de *modèle* permet de contourner le problème. En effet, de par la construction des modèles, une théorie axiomatisée qui possède au minimum un modèle est d'emblée consistante. Un modèle $M =< U, R, a >$ est une structure $S =< U, R >$, où U est un univers du discours ($U \neq \varnothing$) et R un ensemble de relation(s) à l'intérieur de U, à laquelle on ajoute une fonction a qui assigne des valeurs de vérité aux propositions dans U. Le modèle nous assure la consistance puisque par définition si une proposition A est vraie, alors sa négation est fausse. De fait, il est impossible d'avoir une proposition A qui est à la fois vraie et fausse en même temps:

$$\vDash_M A \Leftrightarrow \nvDash_M \neg A$$

Cela dit, la notion de modèle nous amène à faire une distinction entre *vérité* et *validité*. Alors qu'une proposition A est dite *vraie* dans un modèle M si la fonction a lui assigne la valeur de vérité *vrai* (i.e. $\vDash_M A \Leftrightarrow a_M(A) = \mathrm{T}$), une proposition est dite *valide* (logiquement) si et seulement si elle est vraie pour toute interprétation M:

$$\vDash A \Leftrightarrow \forall M, \vDash_M A$$

La stratégie afin de montrer qu'un système formel K est consistant est de montrer que les règles d'inférence et que les axiomes de K préservent la validité (ou la vérité) dans M. Il s'agit là de la preuve d'adéquation du système (*soundness*), à savoir montrer que si A est une conséquence syntaxique de K, alors A est une conséquence sémantique de M:

$$\vdash_K A \Rightarrow \vDash_M A$$

Autrement dit, en montrant que les règles et les axiomes de K préservent la vérité dans une interprétation M, on en vient à montrer

[17] Cf. second théorème d'incomplétude de Kurt Gödel, « On Formally Undecidable Propositions of Principia Mathematica and Related Systems, » dans *From Frege to Gödel: A Source Book dans Mathematical Logic, 1879-1931*, éd. Jean van Heijenoort (Harvard, 1931).

que tous les théorèmes de K sont vrais dans M, et donc que si A est un théorème de K, alors A est vrai dans M. Cela permet de montrer la consistance de K puisque toute contradiction à l'intérieur de ce dernier se retrouverait aussi dans le modèle. En effet, puisque M est un modèle de K, il s'ensuit que si K est inconsistant, alors la contradiction se retrouvera aussi dans M:

$$\vdash_K \bot \Rightarrow \vDash_M \bot$$

Toutefois, par construction de M une telle situation est impossible puisqu'il est impossible d'avoir à la fois A et $\neg A$. vrais dans M. Or, puisque M est un modèle de K et que $\nvDash_M \bot$ il est possible de conclure par *modus tollens* que $\nvdash_K \bot$, et donc que K est consistant.[18]

En vertu de la distinction entre *vérité* et *validité*, on retrouve trois types de propositions à l'intérieur d'un modèle, en l'occurrence les propositions tautologiques, contradictoires et *satisfaisables*, c'est-à-dire qui peuvent être vraies ou fausses relativement à une interprétation. Or, l'intérêt des propositions satisfaisables d'un point de vue formel est que l'ajout de celles-ci à certains systèmes de base permet de créer des théories axiomatisées. En effet, l'axiomatisation consiste à prendre un système initiale K_0, comme par exemple la logique du premier ordre, et à ajouter des axiomes (propositions satisfaisables) de façon à préciser l'ensemble des structures dans lesquelles le système formel pourra être interprété. Par exemple, la logique du premier ordre à laquelle on ajoute l'axiome

(A1) $$(\forall x)(R_1(x, x))$$

ne peut référer qu'à une structure à l'intérieur de laquelle se trouve une relation réflexive. De fait, en ajoutant des propositions

[18] Un point important à souligner est que même s'il est possible de statuer sur le fait que les contradictions ne sont pas des théorèmes des systèmes formels qui admettent des modèles, cela ne donne pas pour autant une procédure de décidabilité pour les non théorèmes. Autrement dit, ce n'est pas parce qu'un système possède un modèle qu'il est décidable. Le modèle permet de montrer que les contradictions ne sont pas des théorèmes, mais cela ne permet pas de décider pour toute proposition s'il s'agit d'un théorème ou non. La logique du premier ordre, par exemple, possède un modèle (plusieurs modèles) sans pour autant être décidable. Cela dit, une *théorie* de premier ordre peut néanmoins être décidable.

satisfaisables à titre d'axiomes à un système K_0 on précise l'ensemble des domaines et des relations auxquels pourront référer les symboles du système formel.

L'axiomatisation consiste à prendre un système de base, souvent construit à partir de la logique propositionnelle, qui rendra compte de certains schémas d'inférences valides. En ajoutant des énoncés satisfaisables comme axiomes à K_0, on créera des extensions du système de base, dont les théorèmes seront les conséquences des axiomes à l'intérieur du système initial. Autrement dit, si l'on construit $K_1 = K_0 \cup \Gamma$, où Γ est un ensemble d'axiomes, alors tous les théorèmes de K_1 seront des conséquences de Γ dans K_0:

$$\vdash_{K_1} A \Leftrightarrow \ \Gamma \vdash_{K_0} A$$

Par ailleurs, la relation entre K_0 et K_1 est telle que

$$K_0 \subset K_1$$

c'est-à-dire que K_1 est une extension de K_0 et donc tout théorème du système de base sera aussi un théorème de K_1. Toutefois, l'inverse n'est pas vrai puisque K_1 aura des théorèmes qui ne sont que des propositions satisfaisables pour K_0.

Outre la relation syntaxique entre les théories, on trouve aussi certaines relations sémantiques entre les modèles. Dans la plupart des cas, le système initial aura un modèle standard qui rend compte de la validité propositionnelle (de la relation de conséquence sémantique) à l'intérieur de celui-ci. Ce modèle standard ne sera toutefois pas nécessairement canonique, c'est-à-dire que tous les modèles d'une théorie ne seront pas nécessairement isomorphes. Néanmoins, considérant que le modèle standard vise à rendre compte de la validité propositionnelle d'une théorie, il est souvent possible de le modifier, c'est-à-dire de spécifier le domaine et les relations de *U* afin d'avoir une définition plus précise de la vérité dans ce modèle.

Prenons un exemple pour illustrer ce point. Dans le cas de la logique du premier ordre, il y a un modèle M_0 qui rend compte seulement de la validité propositionnelle. Puisque cette dernière est complète (au sens large), l'ensemble des théorèmes de la logique du premier ordre correspond exactement aux propositions valides dans M_0 :

(complétude) $\vdash_{K_0} A \Leftrightarrow \ \vDash_{M_0} A$

Comme nous l'avons vu, il est possible d'ajouter des axiomes à K_0, lesquels sont satisfaisables dans M_0, de façon à obtenir une extension qui exprime des relations plus précises. Dans le cas où $K_1 = K_0 \cup \Gamma$, nous avons le résultat suivant, à savoir que si A est un théorème de K_1, alors A est une conséquence sémantique de Γ dans K_0 (en vertu de la complétude du calcul des prédicats) :

$$\Gamma \vDash_{M_0} A$$

Cela dit, si K_1 est consistant, alors le système aura d'autres modèles dans lesquels ses théorèmes seront vrais. De fait, K_1 aura un modèle M_1, ce qui n'implique toutefois pas que K_1 soit complet par rapport à M_1. Tout ce qui est vrai dans M_1 sera une conséquence sémantique de Γ dans M_0 :

$$\vDash_{M_1} A \Leftrightarrow \ \Gamma \vDash_{M_0} A$$

Nous avons pris un exemple où le modèle de base conserve la validité propositionnelle. Or, les mêmes remarques s'appliquent lorsque le modèle est plus restreint, et donc nous aurions pu prendre deux modèles arbitraires M_i et M_j où l'on aurait parlé de vérité plutôt que de validité propositionnelle. À partir du moment où un modèle M_j se contruit à partir d'un modèle M_i, auquel on ajoute certaines relations ou que l'on spécifie le domaine, on obtient une relation sémantique entre les deux modèles puisque ceux-ci partagent la même structure de base. Cela se comprend en fonction du fait que les modèles partageront une partie du domaine ou des relations.

En bref, les relations entre les modèles et les systèmes formels sont tributaires du fait qu'il est possible d'avoir des extensions construites à partir d'un modèle (ou d'un système) initial plus général, ce qui rend possible les *extensions syntaxiques* et les *extensions sémantiques*. En vertu de ces extensions et dépendamment de l'adéquation et de la complétude des théories, il devient possible de faire une multitude de liens entre les conséquences du système initial et celles de l'extension. Notamment, ces extensions permettent d'étudier la dépendance entre les propositions, la dépendance entre les théories et permettent aussi de délimiter le domaine auquel peut s'appliquer une théorie.

Un point intéressant de l'axiomatisation est qu'elle permet d'étudier les relations de dépendance entre les propositions, et de façon plus générale entre les théories. En effet, à partir du moment où une théorie K_1 peut être axiomatisée dans le cadre d'un système initial K_0 (où $K_1 = K_0 \cup \Gamma$ et $\Gamma = \{A_1, \ldots, A_n\}$. un ensemble quelconque d'axiomes), il devient possible d'étudier les relations de dépendance entre les axiomes (ou de façon plus générale, les relations entre les énoncés satisfaisables dans M_0 modèle de K_0). Par exemple, en vertu des extensions sémantiques et syntaxiques, il devient possible d'étudier la relation de dépendance entre $\Gamma' = \{A_1, \ldots, A_{n-1}\}$.. et A_n à l'intérieur du système formel K_0. Puisque K_0 possède un modèle M_0, il suffit de trouver un modèle M_1 extension de M_0 qui satisfait l'énoncé $\neg(\Gamma' \supset A_n)$ (ou à l'inverse, $\neg(A_n \supset \Gamma')$). Dans une telle situation, c'est-à-dire où $\vDash_{M_1} \Gamma'$ et $\nvDash_{M_1} A_n$ (ou l'inverse), il est possible de conclure que Γ' et A_n sont indépendants dans K_0:

(adéquation)	$\Gamma' \vdash_{K_0} A_n \Rightarrow \Gamma' \vDash_{M_0} A_n$
(ext. sém.)	$\Gamma' \vDash_{M_0} A_n \Rightarrow \Gamma' \vDash_{M_1} A_n$.
($\vDash_{M_1}$)	$\nvDash_{M_1} \Gamma'$ ou $\vDash_{M_1} A_n$.

Or, s'il est possible de trouver un modèle M_1 où à la fois Γ' est vrai et A_n est faux, il est possible de conclure par *modus tollens* que A_n n'est pas une conséquence de Γ' dans K_0, et donc que ceux-ci sont indépendants. Plus précisément, l'exemple montre que la proposition A_n est indépendante de l'ensemble de propositions Γ'. En plus de permettre d'étudier les relations entre les propositions, la méthode axiomatique permet aussi d'étudier la relation de dépendance entre les ensembles de propositions. En ce sens, si deux théories différentes sont axiomatisées à partir d'un même système formel initial, il devient possible d'étudier la dépendance entre les théories, c'est-à-dire observer si l'une est la conséquence de l'autre, ou encore voir quels axiomes doivent être ajoutés à l'une des théories afin que les deux soient équivalentes (ou que l'une soit la conséquence de l'autre). Autrement dit, l'axiomatisation permet de hiérarchiser les théories commensurables, c'est-à-dire celles qui ont une structure minimale en commun, puisqu'elle permet de montrer que certaines théories sont plus fortes, plus faibles, équivalentes ou simplement indépendantes face à d'autres théories. L'étude de la dépendance entre les propositions permet de vérifier que les axiomes d'une théorie sont indépendants (ou non) des axiomes d'une autre théorie.

D'un point de vue théorique, l'étude de la dépendance entre les propositions (ou ensembles de propositions) est fort pertinente. En effet, cela permet entre autres de répondre à la deuxième partie du présupposé de Hilbert, à savoir que tout problème possède une solution. Lorsqu'une théorie K_1 est axiomatisée, il devient possible de prouver que certaines conclusions ne peuvent pas être obtenues à partir d'un ensemble d'hypothèses dans le cadre théorique de K_1. Tel que susmentionné, il suffit de trouver un modèle de K_0 (K_1 extension de K_0) où les hypothèses sont vraies mais la conclusion fausse. De fait, la méthode axiomatique fournit un outil puissant d'un point de vue conceptuel puisqu'elle permet de montrer que certaines conclusions ne peuvent pas être obtenues dans certains cadres théoriques. Ainsi, la réorientation du problème devient possible, c'est-à-dire que l'on peut tenter de trouver d'autres hypothèses, voire d'autres théories, qui permettront d'obtenir la conclusion désirée. En ce sens, la méthode axiomatique fournit une manière d'observer avec précision la portée de nos théories en permettant de montrer que certains raisonnements nécessitent des cadres théoriques précis.

Outre les biens faits théoriques que la méthode axiomatique apporte au niveau de l'étude des relations entre les théories et relativement à l'impossibilité de certains raisonnements à l'intérieur d'un cadre théorique donné, la formalisation permet aussi de délimiter l'objet de la théorie. En effet, étudier les modèles possibles d'une théorie équivaut à délimiter le champ à l'intérieur duquel celle-ci peut être appliquée. Une fois axiomatisée, différents modèles de la théorie peuvent être construits. Or, en construisant différents modèles, on change le domaine et on fait varier l'ensemble des relations qui s'appliquent à l'intérieur de l'univers du discours. Ce faisant, il est possible de statuer quant à savoir quels sont les modèles d'une théorie et quels sont ceux qui ne le sont pas. Autrement dit, on parvient à déterminer une classe de domaines et de relations dans lesquels la théorie est consistante. En ce sens, l'étude des modèles d'un système formel permet de délimiter son objet, c'est-à-dire de préciser les interprétations que l'on peut faire du système. Considérant que les symboles n'ont pas d'interprétation fixe, il est primordial de vérifier dans quelles conditions le système formel est consistant, et par le fait même déterminer quelles sont les interprétations que l'on ne peut pas faire du système. Par exemple, le système construit à partir de la logique du premier ordre à laquelle on ajoute l'axiome (A1) ne pourrait pas être interprété dans un modèle où $R_1(x,y) = x > y$ puisque la relation *être plus grand que* n'est pas réflexive. Les modèles d'un tel système formel devront donc nécessairement mettre en jeu une relation R_1 réflexive. Bref, l'axiomatisation permet de délimiter l'objet d'une théorie dans la mesure où un ensemble d'axiomes restreint la classe des modèles à l'intérieur desquels une théorie peut s'interpréter. Cela n'implique toutefois pas que le système formel engendre sa propre interprétation. Étant donné que les axiomes expriment les propriétés des relations, il s'ensuit qu'un ensemble d'axiomes a une incidence directe sur la classe des relations auxquelles le système peut référer.

4. Axiomatisation et physique

La méthode axiomatique, qui consiste à formaliser les relations qui se trouvent entre les énoncés et les concepts d'une théorie, est apparue en vue de fournir des fondements solides à la connaissance mathématique. Comme nous l'avons vu, cette méthode détermine les schémas d'inférences valides pour une théorie donnée et ainsi assure la nécessité des raisonnements. Or, outre le fait que cette méthode ait été développée en vue d'assurer la pérennité de la connaissance mathématique, il n'en demeure pas moins que Hilbert voyait là aussi un outil pouvant être appliqué à la physique et ainsi solidifier les bases épistémiques de cette discipline. Après avoir vu brièvement l'émergence de la méthode axiomatique ainsi que ses caractérisations d'un point de vue plus moderne, il nous reste à mettre en lumière le rôle que Hilbert lui accordait en physique. En plus de mettre en évidence la conception que se faisait Hilbert de l'axiomatisation de la physique, l'objectif est de montrer son intérêt théorique mais surtout les limites épistémiques qu'elle permet de déterminer. La présente section n'abordera pas l'axiomatisation de la physique en ce qui concerne son formalisme mais portera plutôt sur l'importance accordée par Hilbert au 6e problème, à savoir l'axiomatisation de toute la physique. Il s'agira de l'épistémologie sous-jacente à l'approche de Hilbert et de l'intérêt qu'il voyait dans le développement d'un tel programme.[19]

Parmi la fameuse liste des 23 problèmes de Hilbert se trouve en sixième position l'axiomatisation de toute la physique. Considérant le succès de l'application de la méthode axiomatique en géométrie, Hilbert y voyait là une possibilité d'adopter le même *point de vue axiomatique*[20], selon lequel la structure logique d'un discours peut s'exprimer par un ensemble d'axiomes, à la physique.[21] En effet, puisque, au même titre qu'en géométrie, la connaissance physique provient d'une intuition spatiale, le point de vue axiomatique, qui consiste en l'axiomatisation des relations que l'on perçoit[22], peut être appliqué en physique vu son succès en géométrie. Cela dit, l'idée derrière l'axiomatisation de la physique était la même que par rapport à la connaissance mathématique, c'est-à-dire que le but était de fournir des fondements solides à la connaissance physique.[23]

L'objectif de Hilbert en ce qui concerne l'axiomatisation de la physique était le même que pour toute application de la méthode axiomatique, à savoir trouver le plus petit ensemble (consistant) d'axiomes qui permet de dériver par règles d'inférence la totalité de la connaissance physique. On retrouve dans les travaux de Hilbert (1915,1917) sur les fondements de la physique une application directe de la méthode axiomatique. En effet, l'auteur se sert de la méthode axiomatique afin d'étudier la dépendance entre les propositions,

---

[19] Pour un examen détaillé du formalisme de Hilbert en physique, voir Leo Corry, *David Hilbert and the Axiomatization of Physics (1898-1918)* (Kluwer, 2004).

[20] Ulrich Majer, « Hilbert's Axiomatic Approach to the Foundations of Science - a Failed Research Program? » dans *Interactions: Mathematics, Physics and Philosophy, 1860-1930*, éd. V. F. Hendricks J. Lützen, K.F. Jorgensen et S.S. Pedersen (Springer, 2006), p. 158.

[21] *Ibid.*, p. 164

[22] Dans ce cas-ci, des relations physiques que l'on perçoit dans le monde.

[23] Le succès de la méthode axiomatique en géométrie réside entre autres dans le fait que son interprétation arithmétique est consistante. Autrement dit, la géométrie est solidement fondée d'un point de vue épistémologique puisqu'il est possible d'en faire une description mathématique. L'idée sera la même pour l'axiomatisation de la physique.

comme par exemple dans le cas où il cherche à montrer que les phénomènes électrodynamiques sont l'effet de la gravitation[24] et que les équations de Maxwell sont la conséquence des équations gravitationnelles[25]. En ce sens, Hilbert voyait dans l'axiomatisation de la physique la possibilité de mettre à l'œuvre la méthode axiomatique et ainsi étudier la structure logique interne aux théories physiques. Or, outre l'intérêt par rapport aux applications susmentionnées de la méthode axiomatique, Hilbert voyait dans l'axiomatisation de la physique la possibilité d'établir l'unité de la connaissance. Effectivement, Hilbert cherchait à fournir une théorie unifiée qui, par le biais des équations du champ (*world equations*), permettrait de déduire l'ensemble de la connaissance physique. En d'autres termes, Hilbert voyait les différentes théories physiques comme étant des branches de la mécanique, à savoir que l'ensemble de la connaissance physique pouvait se réduire aux axiomes de la mécanique[26] :

> As one can see, the few simple assumptions expressed in axioms I and II [the world equations] suffice with appropriate interpretation to establish the theory: through it not only are our views of space, time, and motion fundamentally reshaped in the sense explained by Einstein, but I am also convinced that through the basic equations established here the most intimate, presently hidden processes in the interior of the atom will receive an explanation, and in particular that generally a reduction of all physical constants to mathematical constants must be possible — even as in the overall view thereby the possibility approaches that physics in principle becomes a

[24] David Hilbert, « The Foundations of Physics (First Communication), » dans *The Genesis of General Relativity*, éd. Jürgen Renn (Springer (2007), 1915), p. 1005.

[25] *Ibid.*, p. 1012

[26] Jürgen Renn et John Stachel, « Hilbert's Foundations of Physics: From a Theory of Everything to a Constituent of General Relativity, » dans *The Genesis of General Relativity*, éd. Jürgen Renn (Springer, 2007), p. 864.

> science of the type of geometry: surely the highest glory of the axiomatic method...[27]

Ce passage, qui conclut la première communication de Hilbert sur les fondements de la physique, est extrêmement révélateur. On y voit non seulement l'optimisme de Hilbert quant à la possibilité d'obtenir une théorie unifiée de la physique qui, à partir des équations du champ, permettrait d'expliquer jusqu'aux comportements des atomes, mais on y voit aussi des indices relativement à son attitude épistémique. L'ampleur du sixième problème y est mise en évidence dans la mesure où l'axiomatisation de la physique vise à assurer une certaine forme d'objectivité de par la description mathématique du monde, mais aussi prétend à pouvoir réduire l'ensemble de la connaissance physique à un groupe restreint de relations. Dans la première partie du passage, Hilbert mentionne que seuls deux axiomes, lorsque interprétés correctement, permettent de rendre compte de la conception einsteinienne de l'espace-temps, alors que vers la fin il est question de la possibilité de réduire toutes les constantes physiques à des constantes mathématiques. Or, ces deux passages marquent clairement l'importance épistémologique d'unifier la connaissance physique et de pouvoir l'interpréter mathématiquement de façon à pouvoir en solidifier les fondements. Cela dit, deux questions importantes méritent d'être soulevées concernant le statut de la connaissance, à savoir (1) *quelles sont les relations qui se trouvent entre le scientifique, la théorie et le monde ?* et (2) *quel est le statut de la connaissance qui résulte de l'application de la méthode axiomatique ?*

D'emblée, la méthode axiomatique ne consiste pas en une activité *a priori* à partir de laquelle il serait possible de dériver un éventail de connaissances. Au contraire, la méthode axiomatique ne sert pas à *créer* des théories mais se veut plutôt un outil qui permet d'examiner la structure logique des théories existantes.

> Hilbert conceived the axiomatization of physics not as a definite foundation that has to *precede* empirical research and theory formation, but as a *post-hoc* reflection on the results of such investigations with the aim of clarifying the

[27] Hilbert (1915), *Op. cit.* note 24, p. 1015

> logical and epistemological structure of the assumptions, definitions, etc., on which they are built.[28]

Dans le cas de la physique, l'idée était de formuler des axiomes qui soient en mesure de rendre compte de la connaissance physique actuelle.[29] Cette conception de la méthode axiomatique se comprend en fonction de la réponse que Hilbert fait à la question concernant la source de la connaissance. Même si la connaissance physique est *a priori* dans la mesure où elle provient de l'activité de la raison, il n'en demeure pas moins que la source de la connaissance est en grande partie *empirique*.[30] Or, cela n'implique pas que les théories scientifiques ou que les axiomes soient *dérivés* ou *inférés* du monde empirique. Au contraire, l'intuition spatiale fournit à la raison le matériau à partir duquel il est possible de construire la connaissance. En ce sens, la méthode axiomatique à elle seule ne peut pas produire de connaissance du monde empirique. Il s'agit plutôt d'un outil qui, une fois une théorie construite, permet d'étudier sa légitimité.

Toutefois, considérant que la source ultime de la connaissance physique est le monde empirique, ne serait-il pas juste d'affirmer que nos théories scientifiques sont conformes à la réalité ? En d'autres termes, étant donné que les axiomes expriment des relations perçues par notre intuition spatiale, ne serait-il pas juste de conclure que nos systèmes axiomatiques correspondent à la structure physique du monde ? En un mot, la réponse est *non.* Comme l'a judicieusement souligné Hempel, l'activité scientifique ne consiste pas à *inférer* des théories de l'expérience empirique mais plutôt à les inventer afin d'en rendre raison. Si les théories scientifiques étaient directement inférées du monde empirique, comment serait-il possible de se tromper ? Au même titre que l'homme de science n'a pas de lien direct avec le monde empirique, les théories et relations exprimées par les axiomes ne correspondent pas à la structure physique du monde mais plutôt à la structure de la compréhension que nous avons de celui-ci.

---

[28] Renn et Stachel (2007), *Op. cit.* note 26, p. 863

[29] *Ibid.*, pp. 864-5

[30] Ulrich Majer et Tilman Sauer, « Intuition and the Axiomatic Method dans Hilbert's Foundations of Physics, » dans *Intuition and the Axiomatic Method*, éd. E. Carson et R. Huber (Springer, 2006), p. 222.

D'une part, il y a toujours une certaine prégnance théorique au niveau de la recherche scientifique, c'est-à-dire que toute recherche scientifique est guidée par une hypothèse ou par une théorie donnée. En ce sens, l'homme de science présuppose que le monde empirique possède une certaine structure et tente par la suite de confirmer cette intuition. De fait, considérant la place de l'hypothèse dans l'activité scientifique, il n'y a pas moyen de savoir si celle-ci correspond exactement au monde empirique. Au mieux, une hypothèse ou une relation exprimée par un axiome peut être confirmée. Dans le cas où les données empiriques contredisent une hypothèse, il devient possible d'affirmer par *modus tollens* que celle-ci est fausse. Toutefois, ce n'est pas parce qu'une relation exprimée par un axiome est confirmée par l'expérience empirique que celle-ci est vraie. Il s'agirait là d'une erreur de raisonnement qui consiste à affirmer le conséquent. La relation entre la théorie scientifique et le monde ne consiste donc pas en une correspondance exacte mais plutôt en une approximation de ce que *pourrait être* la réalité.

Par ailleurs, l'homme de science n'a aucun accès direct ou privilégié aux phénomènes.[31] En effet, toute connaissance empirique est médiatisée par nos sens et par la conception que nous avons du monde. Or, nos sens sont limités et rien ne nous garantit que notre compréhension du monde serait la même si nous avions d'autres facultés sensorielles plus ou moins développées. Le point soulevé ici est que la connaissance du monde physique ne porte pas sur sa structure en tant que telle mais correspond plutôt au cadre à l'intérieur duquel nous percevons la réalité. Dès lors, il est vain de croire que les axiomes physiques, lesquels proviennent de notre intuition spatiale, expriment la structure du monde empirique. Les théories scientifiques reflètent la compréhension que nous avons de la réalité : les structures ne sont pas dans le monde mais bien dans notre conception du monde.

Du point de vue axiomatique, notre connaissance se traduit par les différents modèles et interprétations que nous construisons. Cependant, un système axiomatique possède une infinité de modèles. Or, puisque ni le monde empirique ni le système formel n'engendrent

[31] Yvon Gauthier, « The Use of the Axiomatic Method in Quantum Physics, » *Philosophy of science* 38 (1971), p. 432.

eux-mêmes leur propre interprétation, comment pourrions-nous affirmer avec certitude que, parmi l'infinité de modèles, il y a un et un seul modèle *vrai*, c'est-à-dire une seule interprétation exacte de la réalité ? Le fait est que certains modèles seront privilégiés par rapport à d'autres en vertu de la validation empirique qu'ils auront reçue. Toutefois, en aucun cas nous ne pourrons conclure qu'il s'agit *du* vrai modèle qui correspond exactement au monde empirique puisque rien logiquement ne nous permet de conclure une telle chose.

Cela dit, une telle conception de la connaissance n'implique pas que l'objectivité soit impossible en science. Au contraire, il y a certains critères d'objectivité qui permettent de justifier qu'une interprétation soit privilégiée plutôt qu'une autre. Parmi ces critères se trouvent entre autres ceux de testabilité, comme la variation des conditions d'une expérience, le nombre élevé de résultats significatifs qui confirment l'hypothèse, la possibilité de reproduire l'expérience, etc. Mais outre les critères d'objectivité relatifs aux conditions dans lesquelles sont faites les expériences, on trouve chez Hilbert le désir d'obtenir une certaine objectivité par le biais de l'application de la méthode axiomatique. En effet, dans une perspective épistémologique, la méthode axiomatique appliquée à la physique est un moyen de graduellement faire le passage de la connaissance empirique subjective à l'objective.

Pour formuler cela dans les termes de Hilbert, il est nécessaire de faire la différence entre le *caractère empirique* de la connaissance et les questions qui concernent sa *validité exacte*.[32] Le développement d'une théorie se veut la construction mentale d'un *cadre de concepts* alors que l'application de ce cadre aux phénomènes empiriques se règle par certains critères d'objectivité en science (e.g., les critères de testabilité). Le caractère empirique des axiomes provient du fait que ceux-ci peuvent être validés ou invalidés par l'expérience.[33] Toutefois, malgré que les théories scientifiques aient un caractère empirique, c'est-à-dire qu'elles soient testables et visent à expliquer certains phénomènes, il n'en demeure pas moins qu'il nous est impossible de statuer quant à savoir si les structures exprimées par nos systèmes axiomatiques correspondent exactement à celles des phénomènes. De

[32] Majer et Sauer (2006), *Op. cit.* note 30, p. 224
[33] *Ibid.*, p. 225

fait, l'activité scientifique requiert des critères d'objectivité qui permettront de justifier l'acceptation d'une théorie plutôt qu'une autre.

En plus de ces critères d'objectivité, la méthode axiomatique est un outil qui permet d'atteindre une certaine objectivité en science dans la mesure où elle permet de dégager les présupposés subjectifs implicites à nos théories scientifiques. L'épistémologie véhiculée par le programme de Hilbert est une forme d'*épistémologie récursive*, où l'évolution de la connaissance scientifique se traduit par une sorte de *purification* des modèles, c'est-à-dire par le découvrement des présupposés subjectifs et anthropomorphiques relatifs à notre perception du monde.

> As time proceeds, this part [the *a priori part*] diminishes, because in the process of logical investigations and experimental testing we reveal more and more of these tacit assumptions, in particular those *anthropomorphic* presuppositions of which we were not aware previously, but on which our theories were built. This inquiry is a *recursive* process, in which we proceed in the direction of an objective (i.e., less subjective) knowledge of nature on the basis of what we have achieved so far. It is a slow an presumably unending process.[34]

Voyons un exemple afin d'éclairer ce point. Une façon de le comprendre est de voir pourquoi Hilbert était prêt à rejeter la géométrie euclidienne.[35] Pour Hilbert, même si cette géométrie est conforme à nos intuitions spatiales, il n'en demeure pas moins que rien ne garantit sa *validité exacte*. En fait, l'axiomatisation de la physique permet de dégager le présupposé anthropomorphique euclidien :

> According to my exposition, physics is a four-dimensional pseudo-geometry, whose metric $g_{\mu\nu}$ is connected to the electromagnetic quantities, i.e. to the matter, by the basic equations (4) and (5) of my first communication. With this

[34] *Ibid.*, p. 229
[35] *Ibid.*, p. 228

> understanding, an old geometrical question becomes ripe for solution, namely whether and in what sense Euclidean geometry — about which we know from mathematics only that it is a logical structure free from contradictions — also possesses validity in the real world.[36]

Autrement dit, l'axiomatisation de la physique nous permet de nous dégager de la géométrie euclidienne, qui, provenant de notre intuition spatiale, était jusqu'à présent un présupposé anthropomorphique implicite à notre conception de la structure du monde empirique. Or, en axiomatisant la physique, il devient possible de se dégager de ce présupposé et ainsi atteindre une connaissance plus objective.

> By revoking the Euclidean geometry as a general presupposition of physics, the theory of relativity maintains instead that geometry and physics have identical character and are based as *one* science on a common foundation.[37]

En ce sens, l'axiomatisation de la physique permet d'atteindre une forme d'objectivité dans la mesure où, en se dégageant des présupposés subjectifs, il devient possible d'unifier la connaissance en montrant que la géométrie peut être dérivée des axiomes de la mécanique.[38] L'évolution de la connaissance se fait donc par le dévoilement et le rejet des présupposés subjectifs, ce qui peut se voir dans l'exemple concernant la géométrie euclidienne, afin d'en arriver à une conception plus *objective* de la physique.

---

[36] David Hilbert, « The Foundations of Physics (Second Communication), » dans *The Genesis of General Relativity*, éd. Jürgen Renn (Springer (2007), 1917), p.1025.

[37] Hilbert (1917), *Op. cit.* note 36, p. 1026

[38] Il ne s'agit pas ici de discuter de la question relative au succès ou à l'échec d'une telle entreprise. À ce jour, il n'y a pas de théorie unifiée de la physique et la question eu égard à la possibilité même d'une telle entreprise reste ouverte. Quoi qu'il en soit, l'objectif ici n'est pas de discuter de la réalisation d'un tel projet mais plutôt d'expliciter l'épistémologie qui, chez Hilbert, allait de pair avec la méthode axiomatique et l'axiomatisation de la physique.

Finalement, la méthode axiomatique permet aussi d'atteindre l'objectivité en science puisqu'elle permet la description mathématique du monde.[39] En effet, en formalisant une théorie, il devient possible d'en faire une interprétation mathématique, comme dans le cas de l'interprétation arithmétique de la géométrie, et ainsi lui fournir une justification plus solide, quoi que cela ne suffise pas à garantir son objectivité. La description mathématique du monde empirique, qui se fait par le biais de l'axiomatisation des théories physiques, est une autre facette de l'objectivité scientifique.

Évidemment, l'intérêt d'une telle méthode n'est pas pratique mais bien théorique et épistémologique. Considérant que la méthode axiomatique ne permet pas de produire de connaissances empiriques, il va de soi que les conséquences positives qu'elle peut engendrer ne seront pas d'une telle nature.[40] Néanmoins, l'axiomatisation demeure un outil puissant d'un point de vue théorique. Au niveau épistémologique, la méthode axiomatique permet de sécuriser les fondements d'une discipline en assurant que la théorie et ses postulats sont consistants. De plus, une telle entreprise permet le progrès scientifique dans la mesure où cela permet d'identifier certaines lacunes internes aux théories. L'axiomatisation permet d'améliorer notre compréhension des théories physiques puisqu'elle met en lumière les relations de dépendance et d'indépendance qui se trouvent entre les axiomes et théorèmes de celles-ci. Par l'étude de la structure logique d'une théorie, l'axiomatisation permet d'améliorer notre compréhension de cette dernière et ainsi renforcer son statut épistémique.[41]

> ... a *real* progress frequently first becomes possible if the 'hidden' logical gaps and other errors such as inner contradictions have been identified and eliminated.[42]

En ce sens, d'un point de vue théorique, les bienfaits de la méthode axiomatique sont incontestables :

[39] Majer et Sauer (2006), *Op. cit.* note 30, p. 215
[40] Majer (2006), *Op. cit.* note 20, p. 175
[41] *Ibid.*, p. 176
[42] *Id.*

> The axiomatic method is a mean of investigating the logical structure of a theory, the dependence and independence of its sentences, and its deductive completeness and consistency.[43]

En bref, même si la méthode axiomatique ne fournit pas un moyen pour *produire* des connaissances empiriques, il n'en demeure pas moins que ses bienfaits théoriques sont incontestables. L'étude de la dépendance entre les propositions et les théories est très pertinente d'un point de vue épistémologique puisqu'elle permet de délimiter clairement la portée et la force des théories. Dès lors, l'importance accordée par Hilbert au sixième problème prend toute sa signification. En voulant axiomatiser la physique, ce dernier cherchait non seulement à lui assurer une certaine pérennité en la dégageant des présupposés subjectifs implicites à nos théories, mais aussi à solidifier son statut et à unifier la connaissance.

Conclusion

Somme toute, l'application de la méthode axiomatique aux sciences empiriques reste pertinente même si celle-ci a été développée en vue de solidifier la connaissance mathématique. Le point de vue axiomatique permet de tracer certaines limites épistémiques, à savoir que les axiomes ne reflètent pas la structure du monde mais plutôt la compréhension que nous en avons. Dans cette optique, la quête de la connaissance, qu'elle soit scientifique ou autre, prend une toute autre signification. Au lieu de chercher à élucider les *mystères de la réalité*, l'activité scientifique consiste plutôt en l'élaboration d'un édifice cohérent de connaissances, qui portent non pas sur les propriétés intrinsèques du monde mais bien sur l'explication que nous avons de celui-ci. Les systèmes axiomatiques n'expriment pas les relations du monde empirique : il s'agit de la formalisation des relations que l'observateur perçoit dans le monde. Les modèles sont des interprétations du monde empirique, c'est-à-dire des façons de concevoir et de comprendre les phénomènes.

[43] Majer et Sauer (2006), *Op. cit.* note 30, p. 223

Malgré le fait que l'axiomatisation ne permette pas de *produire* des connaissances, il n'en demeure pas moins que cette méthode a une portée épistémique importante, ce qui peut se voir entre autres par le biais de l'examen des critères de scientificité. Ces critères, que l'on peut diviser en deux groupes, en l'occurrence ceux des critères théoriques et pratiques, incluent d'une part la consistance, la limitation de l'objet, l'irréductibilité et l'analyticité, et d'autre part la complétude (couvrir l'ensemble des phénomènes auxquels la théorie s'applique), l'applicabilité, la prédictivité et la falsifiabilité (la prédiction doit être vérifiable). Or, lorsque l'on examine la répercussion de la méthode axiomatique sur les critères théoriques de scientificité, il devient difficile de ne pas reconnaître son importance. D'emblée, la consistance d'une théorie est l'un des critères de rationalité des plus importants et la méthode axiomatique vise précisément à établir ce point. La consistance étant une affaire de logique[44], il est évident que la méthode axiomatique est pertinente du point de vue scientifique. Par ailleurs, l'étude des différents modèles d'un système formel permet de délimiter l'objet d'une théorie en montrant clairement quels sont les domaines auxquels une théorie axiomatisée peut s'appliquer. De plus, l'examen qui a été fait à la section 3 concernant les relations de dépendance entre les propositions et les théories fournit un moyen d'étudier l'irréductibilité d'une théorie. En effet, grâce à la méthode axiomatique, il est possible de s'assurer qu'une théorie n'est pas réductible à d'autres théories existantes. Finalement, et c'est là un des points les plus évidents, l'axiomatisation d'une théorie répond directement au critère d'analyticité, à savoir qu'une théorie scientifique doit pouvoir s'exprimer par le biais d'un système formel. Ayant cela en tête, il devient clair que la méthode axiomatique, lorsque appliquée aux sciences empiriques ou à tout autre domaine de la connaissance, est un outil puissant qui permet de solidifier l'édifice de la connaissance et d'éclairer ses limites.

---

[44] Yvon Gauthier, « Hilbert's Idea of a Physical Axiomatics: The Analytical Apparatus of Quantum Mechanics, » *Journal of Physical Mathemathics* 2 (2010), p. 1.

Références

CORRY, L., 2004. *David Hilbert and the Axiomatization of Physics (1898-1918),* Kluwer.

FREGE, G., 1980. *Philosophical and Mathematical Correspondence*. Édité par Gottfried Gabriel. Oxford: Basil Blackwell.

GAUTHIER, Y., 2010. « Hilbert's Idea of a Physical Axiomatics: The Analytical Apparatus of Quantum Mechanics. » *Journal of Physical* Mathemathics 2.

——— 1971. « The Use of the Axiomatic Method in Quantum Physics. » *Philosophy of science* 38, pp. 429-37.

GÖDEL, K., 1931. « On Formally Undecidable Propositions of Principia Mathematica and Related Systems. » Dans Jean van Heijenoor (ed), 2002. *From Frege to Gödel: A Source Book in Mathematical Logic*, 1879-1931, Harvard University Press.

HILBERT, D., 1923. « Foundations of Mathematics. » Dans William Ewald (éd.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Oxford: Clarendon Press.

———, 1928. « The Foundations of Mathematics. » Dans Jean van Heijenoor (éd.), 2002. *From Frege to Gödel: A Source Book in Mathematical Logic*, *1879-1931*, Harvard University Press.

———, 1915. « The Foundations of Physics (First Communication) » Dans, Jürgen Renn (éd.), 2007, *The Genesis of General Relativity,* Springer.

———, 1917. « The Foundations of Physics (Second Communication) » Dans, Jürgen Renn (éd.), 2007, *The Genesis of General Relativity*, Springer.

———, 1900. « From Mathematical Problems. » Dans William Ewald (ed), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Oxford: Clarendon Press.

———, 1922 « The New Grounding of Mathematics. » Dans William Ewald (ed.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Oxford: Clarendon Press.

———. 1900, « On the Concept of Number. » Dans William Ewald (ed.)*, From Kant to Hilbert: A Source Book in the Foundations of Mathematics, Oxford: Clarendon Press.*

MAJER, U. « Hilbert's Axiomatic Approach to the Foundations of Science - a Failed Research Program? » Dans V. F. Hendricks, J.

Lützen, K.F. Jorgensen et S.S. Pedersen (éd.), 2006. *Interactions: Mathematics, Physics and Philosophy, 1860-1930,* Springer.

MAJER, U et SAUER, T., « Intuition and the Axiomatic Method in Hilbert's Foundations of Physics. » Dans E. Carson et R. Huber (éd.), 2006. *Intuition and the Axiomatic Method*, par. Springer.

RENN, J. et STACHEL, J., 2007. « Hilbert's Foundations of Physics: From a Theory of Everything to a Constituent of General Relativity. » Dans Jürgen Renn (éd.), 2007. *The Genesis of General Relativity*, Springer.

RUSSELL, B., « Letter to Frege. » Dans Jean van Heijenoor (ed), 1999. *From Frege to Gödel: A Source Book in Mathematical Logic*, *1879-1931*, Harvard University Press.

Preuves intuitionnistes touchant la première philosophie*

Joseph Vidal-Rosset
Université de Nancy 2

A tous les cartésiens

Vuillemin a toujours lu Descartes comme un intuitionniste avant la lettre, si l'on entend par « intuitionniste » un mathématicien qui adopte la philosophie de Brouwer et la logique de Heyting. On se propose dans l'Introduction et la première section de cet article de montrer que cette lecture que Vuillemin fait de Descartes est parfaitement justifiée en expliquant pourquoi les Méditations *peuvent être lues comme une application de la logique intuitionniste. La seconde et la troisième section sont respectivement consacrées à l'analyse logique de la preuve du* Cogito *(Méditation seconde) et de la première preuve de l'existence de Dieu (Méditation troisième). On montre que les deux preuves fondamentales des* Méditations métaphysiques *de Descartes sont toutes les deux valides en logique intuitionniste. Du point de vue logique, la première preuve de l'existence de Dieu que donne Descartes pourrait être considérée comme un progrès par rapport à la preuve d'Anselme qui est concluante en logique classique mais qui échoue en logique intuitionniste. Cependant on nuance ce jugement en conclusion en insistant sur le fait que le concept de Dieu sur lequel Descartes fonde sa preuve est indiscutablement un élément réaliste et non constructif.*

---

* Cet article est très différent de l'exposé que j'ai donné à la conférence de Montréal en septembre 2011. Le projet d'écrire ce texte est né après une discussion engagée par mes étudiants qui assistaient aux conférences données par François Lepage à l'Université de Lorraine en février 2012. Merci à mes étudiants et à François Lepage. Je remercie enfin à travers cet article, tous les logiciens et informaticiens, auteurs d'outils précieux pour la vérification des théorèmes, mis à la disposition de tous, sous forme de logiciels libres.

Introduction
De l'intérêt de l'analyse intuitionniste des preuves des *Méditations*

Le projet philosophique de Descartes déclaré dès la première phrase de la première Méditation est d'« établir quelque chose de ferme et de constant dans les sciences », et pour cela, de commencer « tout de nouveau, dès les fondements », en *démontrant*, à partir de la méthode du doute, des vérités absolument certaines pour toute conscience. C'est parce qu'il n'y a pas de théorie de la démonstration sans étude de la logique qu'il est crucial d'analyser la logique des preuves des *Méditations* pour pouvoir les comprendre. Certes, la syllogistique était perçue par Descartes comme une discipline formelle mais vide de contenu intuitif. Il est vrai que le fondement rationnel de nos connaissances *sur le monde* n'est pas à chercher, pour Descartes, du côté de la logique, de laquelle il y a selon lui finalement peu à apprendre, mais du côté des mathématiques, c'est-à-dire de l'arithmétique et de la géométrie. Cependant, pour écarter immédiatement tout malentendu, il faut rappeler que, du point de vue intuitionniste, la logique est une *partie* des mathématiques et que, dès lors qu'elle est corrigée des défauts de la logique classique, elle ne doit être écartée ni des mathématiques, ni de l'entreprise du fondement des sciences. La logique intuitionniste, contrairement à la logique classique est fondée sur une représentation de la progression de la connaissance dans le temps, et c'est cet aspect fondamental qui la rend particulièrement adaptée pour la compréhension des *Méditations* en raison de l'analogie suivante : de la même façon que la méthode du doute permet de comprendre ce qui est métaphysiquement nécessaire, la logique intuitionniste donne aussi les règles générales des preuves mathématiques qui sont recevables d'un point de vue intuitionniste. Cette logique forme donc le « noyau dur » de l'intuitionnisme et occupe une place comparable à celle que Descartes donne aux *Méditations* dans son système. En effet, les *Méditations métaphysiques* fondent à leur tour la science mathématique elle-même sur la preuve du *Cogito* qui conduit par une chaîne de raisons aux preuves de l'existence de Dieu, ce dernier étant le garant de la vérité des idées claires et distinctes que toute conscience peut reconnaître dans les éléments de l'arithmétique et de la géométrie.

J'ai déjà souligné ailleurs [**19**] que Vuillemin entend par « intuitionnisme » tout système de philosophie de la connaissance qui, comme celui de Kant après celui de Descartes, place le Sujet au centre de la constitution de sa description intégrale et systématique de la réalité ; autrement dit, Vuillemin ne restreint pas le sens de ce terme à la philosophie des mathématique qui est née au début du vingtième siècle avec Brouwer. Scrupuleusement attentif à l'histoire de la philosophie, Vuillemin [**24, 25**] a remarqué que la philosophie de Descartes s'est constituée comme celle d'Épicure en privilégiant « les jugements de méthode » ; le doute cartésien étant le type même d'un jugement de méthode, c'est-à-dire un performatif théorique systématiquement employé pour décider du vrai. En considérant Descartes comme un intuitionniste, Vuillemin évidemment invite à comparer la philosophie de Descartes et celle de Brouwer, et il est vrai qu'elles ne sont pas sans points communs.

L'œuvre révolutionnaire de Brouwer en mathématiques est fondée sur l'idée selon laquelle les mathématiques relèvent d'une libre activité de l'esprit indépendante d'un quelconque langage particulier. Dès 1908, Brouwer manifeste une méfiance à l'égard de la logique formelle classique qui s'exprime par ces mots [**4**, pp. 18-19] :

> […] les raisonnements logiques effectués indépendamment de la perception, attendu qu'ils sont les signes de transformations mathématiques à l'intérieur du système mathématique qui régit les perceptions, peuvent déduire, de prémisses scientifiquement admises, des conclusions inacceptables.

Dans ce même article, Brouwer remarque au passage que « la sagesse qui se manifeste dans l'œuvre de Spinoza est ressentie comme totalement indépendante de sa systématisation logique », avant de souligner que les paradoxes logico-mathématiques contemporains sont nés d'un privilège indûment accordé à la logique sur une mathématique vidée à tort de l'intuition originaire du temps. Il poursuit et achève son analyse du rôle de la logique en mathématiques en affirmant que, si le principe du syllogisme et celui de contradiction sont universellement fiables, le principe du tiers exclu en revanche ne l'est plus pour les systèmes infinis.

On sait aujourd'hui comment Heyting et Kolmogorov ont réussi à exprimer une logique, dite « intuitionniste » qui est capable de répondre aux exigences de Brouwer, et comment les modèles et contre-modèles de Kripke donnent une interprétation formelle claire à cette logique. Je me propose dans cet article de montrer que, du point de vue de l'argumentation philosophique, les trois premières Méditations de Descartes illustrent avec force exactement la même position. En 1960, Vuillemin [**22**] avait déjà montré que la *Géométrie* de Descartes repose sur une décision métaphysique qui « insère l'œuvre cartésienne dans la tradition des mathématiques intuitionnistes sinon finitistes ». Mais probablement parce qu'il a toujours été convaincu du bien fondé et de la solidité de la logique classique et qu'il n'a jamais prêté une grande attention au développement technique de la logique intuitionniste, Vuillemin, à ma connaissance, ne semble s'être jamais vraiment posé la question de savoir si la recherché des preuves qui est à l'œuvre dans les *Méditations* s'accorde avec les principes fondamentaux de la philosophie de Brouwer et avec la logique intuitionniste en général. Je vais montrer qu'il en est bien ainsi.

1. Première Méditation
Des choses que l'on ne peut logiquement révoquer

Je développe dans cette section quelques éléments qui justifient encore un peu plus le rapprochement que l'on peut faire entre la philosophie de Descartes et l'intuitionnisme contemporain et je montre pourquoi l'on est fondé à penser, à la lecture de la Première Méditation, que la logique des démonstrations dans les *Méditations* est la logique intuitionniste.

L'entreprise philosophique de Descartes, tout comme celle de Brouwer, repose avant tout sur cette conviction inébranlable : la connaissance rationnelle ne peut être fondée de manière ferme et constante que sur des idées simples, claires et distinctes pour tout sujet pensant. On aurait tort de négliger l'originalité d'une telle position, car elle ne va pas de soi pour nombre de philosophes qui n'accordent aucun privilège au sujet pensant en tant qu'être singulier isolé de l'activité sociale et collective. Pour un Quine comme pour bien d'autres, les vrais doutes sont définis par les problèmes que posent les sciences ; en conséquence de quoi le point de départ de Descartes a, pour tout philosophe « dogmatique » au sens que Vuillemin donne à ce terme, quelque chose d'artificiel et de faussé. Le rejet spinoziste de la méthode du doute cartésien, exprimé par exemple dans la proposition **XLIII** de la seconde partie de l'*Ethique* [**16**] - « Qui a une idée vraie sait en même temps qu'il a une idée vraie et ne peut douter de la vérité de sa connaissance » -met en relief l'originalité de la position cartésienne. Pour Descartes, le fait que je sois capable de parvenir jusqu'au doute métaphysique et hyperbolique, en forgeant la fiction selon laquelle je pourrais être trompé par une puissance infiniment trompeuse, même lorsque je fais l'addition de deux et de trois, atteste de la liberté de la volonté qui est présente dans tout jugement, selon la théorie de la troisième Méditation. Comme le dit très justement Alquié [**1**, p. 177] :

> la marche de la Méditation première relève avant l'heure de la structure ontologique du « je pense » : du sensible, on doute par raison, car l'entendement dépasse et fonde le sensible ; du rationnel, on doute par volonté, car la volonté est infinie, dépasse l'entendement, et constitue le fond le plus authentique de nous-mêmes.

Dans sa note introductive à l'article cité de Brouwer [**4**, p. 16], Largeault écrit que « l'origine de l'intuitionnisme brouwerien est une métaphysique où le vouloir est antérieur à la pensée ». Je laisse aux exégètes de la pensée de Brouwer le soin de vérifier jusqu'à quel point cette remarque de Largeault sur Brouwer est correcte. Mais cette citation de Largeault pourrait aussi caractériser correctement la position de Descartes, à la condition d'être modifiée ainsi : « l'origine de l'intuitionnisme *cartésien* est une métaphysique où la *liberté du vouloir est la condition de possibilité de la méthode du doute* ».[1]

Cependant, si le doute est de toute évidence un acte de la volonté, il n'est pas moins évident qu'il faut, dans le contexte cartésien, en

[1] Il est tout à fait remarquable que la négation ne peut pas être comprise adéquatement de manière intuitionniste si l'on n'admet pas pour la décrire *un élément pré-logique*, qui exprime en réalité l'action, et non la représentation. Mais j'abandonne ici le développement de cette remarque pour rester dans les limites de mon sujet.

concevoir l'usage comme une méthode pour la recherche de la vérité, et non comme le font les sceptiques, comme une fin en soi. Ce point mériterait à peine d'être mentionné si, dans la célèbre et prestigieuse revue *Analysis*, dans un article publié en 1979 [**17**], on ne pouvait pas au sujet de Descartes, lire ce genre de perles :

> In one of the most celebrated philosophical passages, the first of his six *Meditations*, Descartes claims to have demonstrated a sceptical conclusion concerning the senses by arguing that the very sensations which seem to confirm the existence of an "external" world might be occasioned by a dream.

Il est regrettable que, dans cet article, l'auteur prenne la première Méditation de Descartes comme simple prétexte et commette un grossier contresens, par manque de fidélité à *l'ordre des raisons* sur lequel Gueroult [**7, 8**] a insisté à juste titre. Il est vrai que Descartes reprend ici à son compte le doute sceptique au sujet des sens, mais il est inexact d'écrire qu'il prétend que ce doute *démontre* les conclusions sceptiques. L'expression de « Cartesian scepticism » utilisée plus loin dans le même article, est incompatible avec ce texte de la sixième Méditation :

> [...] pouvant user de ma mémoire pour lier et joindre les connaissances présentes aux passées, et de mon entendement qui a déjà découvert toutes les causes de mes erreurs, je ne dois plus craindre désormais qu'il se rencontre de la fausseté dans les choses qui me sont le plus ordinairement représentées par mes sens. Et je dois rejeter tous les doutes de ces jours passés, comme hyperboliques et ridicules, particulièrement cette incertitude si générale touchant le sommeil, que je ne pouvais distinguer de la veille : car à présent j'y rencontre une très notable différence, en ce que notre mémoire ne peut jamais lier et joindre nos songes les uns aux autres et avec toute la suite de notre vie, ainsi qu'elle a de coutume de joindre les choses qui nous arrivent étant éveillés.

Cette citation de la sixième Méditation montre que dans la première Méditation, Descartes *simule* un doute sceptique qui le conduit à reconnaître qu'il ne dispose pas *pour le moment* d'une quelconque preuve de la fiabilité des sens. C'est la raison pour laquelle *les raisons naturelles de douter* ne touchent pas les vérités de l'arithmétique et de la géométrie. En effet, les preuves de ces vérités peuvent être faites par un entendement qui ne s'inquiète pas de la correspondance de ses idées avec les objets du monde extérieur :

> Car, soit que je veille ou que je dorme, deux et trois joints ensemble formeront toujours le nombre de cinq, et le carré n'aura jamais plus de quatre côtés ; et il ne semble pas possible que des vérités si apparentes puissent être soupçonnées d'aucune fausseté ou d'incertitude.

Ce moment de la première Méditation établit que, si le doute *naturel* permet de remettre en cause la correspondance des représentations et des choses existantes représentées, en revanche, comme le souligne avec précision Gueroult [**7**, p. 36], il ne touche pas les idées simples et universelles qui sont les conditions nécessaires de toute représentation possible :

> De ce genre de choses est la nature corporelle en général, et son étendue ; ensemble la figure des choses étendues, leur quantité ou grandeur, et leur nombre ; comme aussi le lieu où elles sont, le temps qui mesure leur durée, et autres semblables.

Ce n'est que lorsque Descartes a épuisé les *raisons naturelles de douter*, qu'il évoque pour finir la possibilité d'un doute *métaphysique*. Le raisonnement de Descartes est le suivant : j'ai en moi l'*opinion* qu'il y a un Dieu dont la puissance est infinie et, bien que l'idée de ce Dieu soit celle d'un Dieu bon incompatible avec l'idée de tromperie volontaire, il n'est pas naturellement douteux que je me trompe parfois, sans que je puisse pour autant affirmer qu'un tel Dieu *veuille* me tromper. On pourrait, poursuit Descartes, être tenté de se ranger du côté des athées plutôt que d'admettre une telle puissance infinie (même si l'idée de tromperie est incompatible avec l'idée de Dieu). Mais alors au sujet des athées s'impose la conclusion suivante :

> Toutefois, de quelque façon qu'ils supposent que je sois parvenu à l'état et à l'être que je possède, soit qu'ils l'attribuent à quelque destin ou fatalité, soit qu'ils le réfèrent au hasard, soit qu'ils veuillent que ce soit par une continuelle suite et liaison des choses, il est certain que, puisque faillir et se tromper est une espèce d'imperfection, d'autant moins puissant sera l'auteur qu'ils attribueront à mon origine, d'autant plus sera-t-il probable que je suis tellement imparfait que je me trompe toujours.

La première Méditation s'achève par la mise en place de la méthode du doute hyperbolique qui consiste à rejeter comme fausse toute idée qui *peut* être douteuse ; autrement dit la méthode réduit le connaissable à l'indubitable. Or, par un raisonnement qui mériterait très certainement d'être analysé de manière plus approfondie[2], Descartes affirme que l'imperfection de ma connaissance a une probabilité d'autant plus grande que l'on suppose une imperfection de la cause de mon être. En conséquence l'athée a encore plus de raisons de douter et ne peut ni échapper au *doute naturel* du sceptique, ni même échapper à l'hypothèse sceptique d'une absence de véracité des idées claires et distinctes, car, comme Descartes le souligne dans les *Sixièmes Réponses aux Objections* (cité par Gueroult [**7**, p. 46]) :

> d'autant moins puissant [l'athée] concevra l'auteur de son être, d'autant plus aura-t-il occasion de douter, si sa nature n'est point tellement imparfaite qu'il se trompe, même dans les choses qui lui semblent très évidentes.

A l'issue de la première Méditation, la méthode est donc fixée : il s'agit de partir d'une idée *métaphysiquement indubitable* pour parvenir à

[2] C'est encore une remarque qu'il sera ici impossible de développer. François Lepage a attiré mon attention sur le fait que la logique intuitionniste peut être mathématiquement définie comme le cadre logique de la théorie des probabilités. Le fait que Descartes conçoive dans la sixième Méditations toutes les vérités qui enveloppent l'existence des corps comme des vérités avec un degré de certitude nécessairement inférieur aux vérités de la métaphysique qui fonde les sciences, confirme l'idée que la bonne logique pour Descartes est cette logique intuitionniste qui ne sera définie que bien plus tard dans l'histoire.

un fondement de la science contemporaine qui soit tout aussi indubitable que cette idée. Cette méthode en réalité ne diffère pas de l'interprétation que l'on donne aujourd'hui de la logique intuitionniste, comme je vais maintenant le montrer.

Traduire formellement la démarche démonstrative des *Méditations* dans la logique intuitionniste n'a d'intérêt que si et seulement si cette traduction ou bien éclaire les *Méditations*, ou bien si les *Méditations* donne une interprétation intuitive claire de la signification des preuves en logique intuitionniste. Ce dernier point peut être assez facilement mettre en évidence en se fondant par exemple sur le formalisme des preuves intuitionniste adopté par Bell *et alii* [**3**].

Remarquons qu'à l'instar des intuitionnistes contemporains, Descartes ne considère une idée comme absolument vraie qu'à la condition que celle-ci soit indubitable ou prouvée. Or du point de vue de la logique intuitionniste, le fait d'être douteux pour une idée ne constitue pas une nouvelle valeur de vérité qui s'ajouterait à celles du vrai et du faux (notées respectivement aujourd'hui $\top$ et $\bot$). Le doute est un *acte* qui dans le cadre cartésien dépend de la volonté et, dans le contexte intuitionniste contemporain est le fait de remarquer que, pour tel énoncé A, on ne possède pas de preuve concluante de la vérité de A. Pour traduire le fait qu'un énoncé A « n'est pas connu comme vrai (*is not known to be true*) » c'est-à-dire, plus simplement et en meilleur français, est *douteux*, Bell *et alii* [**3**, pp. 192-223] le notent par l'apposition d'un point d'interrogation devant la formule A qui fera l'objet du test de validité intuitionniste. L'application des règles intuitionnistes pour la méthode des arbres de Beth (voir Appendice A, pages 30-34) permet de donner une traduction simple de la signification de la méthode de preuve indirecte qui rappelle la même méthode dans le cadre de la logique classique : soit $?(A)$, si le développement arborescent complet de toutes les sous formules de $?(A)$ donnent uniquement des branches qui se ferment sur une contradiction c'est-à-dire sur l'absurde, alors le doute de la validité de A est lui-même absurde et donc A est effectivement prouvable du point de vue intuitionniste, ce qui peut s'écrire ainsi :

$$\vdash_i A \quad \leftrightarrow_{def.} \quad (\alpha \nVdash A) \rightarrow \bot \tag{1}$$

ou, dans la notation de Bell *et al.* :

$$\vdash_i A \quad \leftrightarrow_{def.} \quad (?A) \rightarrow \bot \tag{2}$$

De même en logique intuitionniste, si une formule A n'est pas prouvable, alors il existe un contre modèle de A qui peut être représenté par un arbre de Beth de ?(A) où il existe au moins une branche ouverte, et ce contre-modèle correspond à un contre-modèle de Kripke. On a donc :

$$\nvdash_i A \quad \leftrightarrow_{def.} \quad \alpha \nVdash A \tag{3}$$

ou encore

$$\nvdash_i A \quad \leftrightarrow_{def.} \quad (?A) \nrightarrow \bot \tag{4}$$

Il est remarquable que la recherche des preuves dans les *Méditations* corresponde exactement à la façon dont la recherche de preuve est conçu en logique intuitionniste. On a vu plus haut que l'impuissance à prouver la fiabilité des sens ne devait être conçue ni comme une preuve en faveur du scepticisme, ni comme une réfutation de celui-ci. Autrement dit la formule

$$\nvdash_i A \tag{5}$$

ne doit pas être considéré comme équivalente à

$$\vdash_i \neg A \tag{6}$$

car (5) exprime le fait d'échouer à prouver A, mais n'exclut pas le fait que A puisse être prouvé ou réfuté ultérieurement, autrement dit (5) n'implique pas (6). En revanche (6) exprime la réfutabilité intuitionniste de A, ce qui est une propriété persistante dans le temps, donc (6) implique (5).

Si l'on prête attention à la manière dont Descartes conçoit les preuves dans les *Méditations*, on constate que le fameux *ordre des raisons* correspond à un ordre dans le temps, chacune des six *Méditations* correspondant à une journée d'étude, la seconde commençant précisément par ces mots : « La Méditation que je fis hier m'a rempli l'esprit de tant de doutes, qu'il n'est plus désormais en mon pouvoir de les oublier ». Nombreux sont les lecteurs qui, peu attentifs à l'ordre adopté, n'ont pas compris qu'une thèse n'est assumée par Descartes qu'au moment où il en donne la preuve, mais pas avant. Tel est le cas par exemple de la distinction réelle de l'âme et du corps, qui ne peut pas être assumée avant la troisième Méditation, c'est-à-dire pas avant la démonstration de l'existence d'un Dieu qui soit le garant de la vérité des idées claires et distinctes. Cela signifie que, du point de vue démonstratif adopté par les *Méditations*, la thèse matérialiste selon laquelle l'âme et le corps sont une seule et même chose, reste une thèse douteuse, ni réfutée ni prouvée, jusqu'à ce que la démonstration de sa fausseté soit faite à la sixième Méditation. Cette importance accordé à l'ordre temporel des démonstrations a trouvé son expression contemporaine dans les modèles de Kripke (Annexe A) qui offrent actuellement le modèle le plus simple pour interpréter la logique intuitionniste.

On peut conclure cette section sur un doute légitime et répondre à ce doute. Si la logique dont Descartes fait usage dans les *Méditations* est vraiment la logique intuitionniste, il est légitime de se demander s'il existe dans ce texte un traitement intuitionniste du théorème du tiers exclu. La réponse à cette question est affirmative ; ce texte existe, il se trouve au premier alinéa de la seconde Méditation, où Descartes établit la transition avec ce qui tout ce qui précède. Je souligne dans la citation qui suit ce qui est crucial pour justifier la thèse selon laquelle Descartes fait usage du tiers exclu *via* une interprétation intuitionniste :

> La Méditation que je fis hier m'a rempli l'esprit de tant de doutes, qu'il n'est plus désormais en ma puissance de les oublier. Et cependant je ne vois pas de quelle façon je les pourrai résoudre ; et comme si tout à coup j'étais tombé dans une eau très profonde, je suis tellement surpris, que je ne puis ni assurer mes pieds dans le fond, ni nager pour me soutenir au-dessus. Je m'efforcerai néanmoins, et suivrai derechef la même voie où j'étais entré hier, en m'éloignant de tout ce en quoi je pourrai imaginer le moindre doute, tout de même que si je connaissais que cela fût absolument faux ; et *je continuerai* ***toujours*** *dans ce chemin, jusqu'à ce que j'aie rencontré quelque chose de certain, ou du moins, si je ne puis*

> *autre chose, jusqu'à ce que j'aie appris certainement, qu'il n'y a rien au monde de certain.*

Remarquons que Descartes fait usage d'une formule qui est effectivement celle du tiers exclu, mais qu'il lui donne précisément une interprétation intuitionniste : dans la recherche indéfinie d'une idée certaine, c'est-à-dire prouvable, le doute hyperbolique sera désactivé pour au moins une idée dont la démonstration est indubitable, ou bien avec la démonstration qu'il n'existe aucune idée certaine. La formalisation de la disjonction énoncée par Descartes est bien une instance de la formule du tiers exclu, où x est une variable qui prend ses valeurs dans le domaine des idées du *Cogito* et où Px signifie « x est prouvable » :

$$\exists x Px \lor \forall x \neg Px \tag{7}$$

Si l'on fait usage de la logique classique, il est démontrable que (5) est une tautologie, comme on le montre ici avec la méthode des arbres de Beth pour la logique classique :

Théorème 1.1. —

$$\vdash_c \exists x Px \lor \forall x \neg Px$$

Démonstration. —

1. $\neg(\exists x Px \lor \forall x \neg Px)$ ✔
2. $\neg\exists x Px$ (1)
3. $\neg\forall x \neg Px$ (1) ✔
4. $\neg\neg Pa$ (3)
5. $\neg Pa$ (2)

×

□

Dans le cadre classique, la disjonction formulée par Descartes serait une platitude. Mais l'interprétation intuitionniste de (7) ne fait pas de cette formule une tautologie. Ce que veut dire Descartes peut être clairement représenté par le schéma suivant :

$$\begin{array}{c} \vdots \\ \textit{Dubito} \\ \exists x Px \quad \forall x \neg Px \\ \vdots \qquad\quad \vdots \end{array}$$

On peut démontrer que (7) n'est pas une tautologie du point de vue intuitionniste en utilisant la méthode des arbres exposée par Bell *et alii* [**3**] et reprise dans l'Appendice A de cet article :

Théorème 1.2. —

$$\nvdash_i \exists x Px \vee \forall x \neg Px$$

Démonstration. —

$$\begin{array}{c} 1.\ ?(\exists x Px \vee \forall x \neg Px)\ ✔ \\ 2.\ ?\exists x Px\ _{(1)} \\ 3.\ ?\forall x \neg Px\ _{(1)}\ ✔ \\ 4.\ ?Pa\ _{(2)} \\ \hline 5.\ ?\neg Pb\ _{(3)}\ ✔ \\ \hline \\ 6.\ Pb\ _{(5)} \end{array}$$

□

Si l'on doute de cette démonstration, on peut vérifier son résultat à l'aide d'Imogen, un programme écrit par Sean McLaughlin [**13**] pour tester automatiquement le caractère prouvable des formules en logique intuitionniste du premier ordre[3]. Un des avantages d'Imogen est une syntaxe standard et assez intuitive : les symboles

```
forall , exists , ~ , & , | , => , <=>
```

dénotent respectivement la quantification universelle, la quantification existentielle, la négation, la conjonction, la disjonction, l'implication et l'équivalence. Imogen adopte aussi la convention du langage Prolog où les variables sont symbolisées par les majuscules (ou par un mot commençant par une majuscule) et en utilisant les lettres minuscules

[3] http://www.seanmcl.com/software/imogen/

(ou les mots commençant par des lettres minuscules) pour symboliser les prédicats. On verra plus loin que cette souplesse inspirée de Prolog permet l'écriture de formules qui restent très intuitives et donc très faciles à lire. Voici la copie du résultat du test de (5) avec Imogen (la barre verticale « | » symbolisant le connecteur « ou ») :

```
joseph@joseph-Inspiron-530:~$ imogen prove
'exists X.p(X) | forall X. ~ p(X)'
### Warning: using slow symbols ###
% SZS status CounterSatisfiable

The database is saturated. The formula is false!
```

Enfin, pour donner un sens intuitif à ce résultat, on peut remarquer que cette instance de la formule du tiers exclu correspond exactement à un certain type de contre-exemples au tiers exclu donnés par Brouwer et qualifiés de « faibles » (par opposition aux contre-exemples dits « forts » donnés plus tardivement aussi par Brouwer et qui, comme l'explique van Atten [**18**], sont ceux qui produisent directement une contradiction lorsqu'on substitue dans les formules les concepts intuitionnistes à leurs expressions classiques). Un des contre-exemples faibles au tiers exclu donné par Brouwer est l'énoncé suivant :

> « Il y a un chiffre qui apparaît plus souvent que tous les autres dans le développement décimal de π, ou tous les chiffres du développement décimal de π sont tels qu'il est faux que l'un d'entre eux apparaît plus souvent que les autres »

Il est évident que la formalisation de cet énoncé en calcul des prédicats donne aussi la formule (7). Le contre-modèle qui prouve le théorème 1.2 illustre, de manière schématique mais précise, le refus intuitionniste d'assumer la transcendance implicitement contenue dans le tiers exclu classique. Comme le montre le contre-modèle, il est évidemment incontestable qu'au moment où l'énoncé quelconque Pa est prouvé, le doute au sujet de la vérité de cette instance du tiers exclu devient absurde. Mais, ce que souligne Brouwer dans [**4**], est précisément le fait que le tiers exclu perd son statut de vérité logique dans les systèmes *infinis*, en donnant évidemment à ce mot le sens intuitionniste d'*infini potentiel*.

On comprend donc que, de la même façon qu'il est possible de chercher indéfiniment une preuve de l'existence d'un chiffre qui apparaît plus souvent que les autres dans le développement décimal de π, il est envisagé comme possible par Descartes que le *Cogito* puisse effectivement « se noyer » en s'enfonçant *toujours plus loin* dans une recherche qui ne lui assure ni l'existence d'une seule idée certaine, ni la preuve qu'il n'y a rien de certain. Autrement dit n'exclut pas de rencontrer une situation qui corresponde au second monde décrit plus haut par le contre-modèle et qui se prolongerait indéfiniment :

$$\begin{array}{c} ?(\exists x Px \vee \forall x \neg Px) \checkmark \\ ?\exists x Px \\ ?\forall x \neg Px \checkmark \\ ?Pa \\ \hline ?\neg Pb \checkmark \\ \vdots \end{array}$$

Résumons le rôle que peut jouer la logique intuitionniste comme théorie de la démonstration dans les *Méditations*. Pour établir la certitude absolue de n'importe quelle idée A, doutons de A, si l'opération qui consiste à douter de A, rend ce doute absurde, c'est-à-dire si $?(A)$ conduit à une contradiction, alors A est une idée indubitable c'est-à-dire une vérité certaine ; dans le cas contraire on considérera que A n'est pas prouvé. Les règles logiques utilisées seront celles de la logique intuitionniste, car, comme on vient de le voir, la formule (7) est prouvable avec les règles de la logique classique, alors qu'elle ne l'est pas avec celle de la logique intuitionniste. On voit donc qu'en utilisant une logique plus faible, on doit pouvoir obtenir une certitude plus forte. La méthode logique étant définie, on peut maintenant montrer pourquoi la preuve intuitionniste du *Cogito* n'a rien de trivial.

2. Méditation Seconde De la nature intuitionniste de la preuve du Cogito

Une idée qui résiste au doute hyperbolique, c'est-à-dire à la fiction du mauvais génie, en impliquant l'absurde dès lors qu'elle fait l'objet de ce doute est une idée indubitable ou encore une idée vraie ou démontrable. La formule (1), caractéristique de la logique intuitionniste, schématise la recherche de preuves dans les *Méditations*. On ne perdra pas de temps à faire la recension des philosophes qui ont affirmé que le *Cogito* est une vérité triviale. Un des derniers à l'avoir fait plus ou moins ouvertement est Hintikka [**9, 11**], qui n'a fait qu'aggraver l'erreur de Gassendi en s'appuyant sur la logique libre de Lambert pour contester l'intérêt que Descartes donne de la preuve du *Cogito* de Descartes. Le raisonnement d'Hintikka [**9**, pp. 147-149] consiste simplement à remarquer que le fameux *Cogito, ergo sum* de Descartes a la forme logique suivante :

$$Cogito(\text{ego}) \to \exists x(x = \text{ego}) \tag{8}$$

Hintikka remarque alors que, si l'on raisonnait dans une logique libre, c'est-à-dire, selon la définition que Lambert [**12**, p. 35] donne de la logique libre une logique libre d'assomptions existentielles sur les termes généraux et singuliers, la question de savoir si l'on peut inférer J'existe de Je pense serait alors une question vraiment pertinente. Mais tel n'est pas le cas puisque tout énoncé de la forme

$$B(a) \to \exists x(x = a) \tag{9}$$

repose sur une présupposition existentielle véhiculée par l'usage de la constante a. En conséquence de quoi Hintikka semble considérer que la remarque que Gassendi avait faire à Descartes dans les *Cinquièmes Objections* selon laquelle la conclusion du J'existe aurait pu être tirée du fait que je me promène comme de n'importe quelle autre de mes actions, puisque

$$Ambuli(\text{ego}) \to \exists x(x = \text{ego}) \tag{10}$$

partage avec (8) la même forme logique, à savoir (9).

Or l'analyse d'Hintikka sur ce point, du moins dans la section de cet article, néglige à la fois la signification et la preuve que Descartes apporte pour établir la vérité du *Cogito*. La suite de l'article s'oriente vers une explication du *Cogito* comme étant une sorte de performatif théorique, ce qui explique pour Hintikka [**9, 11**, p. 167] le privilège que Descartes accorde au verbe *cogitare*. Néanmoins, la critique développée dans la section 4 de l'article subsiste : la présupposition existentielle d'un énoncé comme (8) n'est pas contestable et rend la preuve de Descartes, aux yeux d'Hintikka, sans intérêt logique.

Pariente est très probablement au fait de la critique d'Hintikka et on peut supposer qu'il songe à son insuffisance lorsqu'il écrit [**15, 14**, p. 39] :

> ce qui est suffisant pour la logique ne l'est pas pour Descartes, en ce point de son argumentation. Plus exactement peut-être, Descartes impose ici une condition qui renforce l'exigence de la logique, en ceci qu'il reconnaît comme vrai que l'indubitablement vrai.

Ici « indubitablement vrai » signifie sans doute « effectivement démontrable ». La réduction du vrai au domaine du démontrable est l'exigence de la logique intuitionniste dont le premier principe théorique est le refus d'assumer l'universalité du principe de bivalence. Il est donc évidemment impropre de parler comme le fait ici Pariente de « *la* logique », comme si l'on pouvait faire abstraction de la différence sémantique fondamentale entre logique classique et logique intuitionniste.

Ce point établi, je crois cependant que, dès lors que l'on raisonne à l'aide la logique intuitionniste, il est possible de démontrer qu'Hintikka comme Pariente accordent tous deux une importance exagérée à la nécessité d'expliquer l'usage de la première personne dans les *Méditations*. Le problème de Descartes n'est pas de prouver l'existence de l'*ego*, mais de trouver une pensée indubitable, c'est-à-dire une pensée qui résiste à la puissance du grand trompeur. En effet, aussi surprenant que cela puisse paraître, il n'y a aucune difficulté à traduire la démonstration de Descartes à la troisième personne, sans rien lui faire perdre de son caractère apodictique. On remarquera que dans la démonstration qui suite, la référence à la première personne du singulier n'est pas requise, mais que l'on peut faire usage de cette première personne pour rendre plus intuitif le résultat.

Supposons donc qu'un sujet décide de considérer toutes ses représentations comme douteuses et forge pour cela la fiction d'un être qui la puissance de le tromper quel que soit le contenu et la nature de ses pensées. Un tel sujet peut-il malgré tout être certain d'au moins une pensée ? La réponse de Descartes à cette question est qu'un tel sujet peut être absolument certain qu'il pense, non seulement parce qu'il est impossible de douter que l'on pense au moment où l'on doute, puisque le doute est une pensée, mais parce que l'acte de tromperie du mauvais génie à l'égard du sujet doutant implique aussi que ce sujet est trompé seulement si ce sujet pense, en vertu de la signification qui est donnée dans ce contexte à cette fiction du mauvais génie. En effet la tromperie est supposée s'exercer sur *les représentations* ou sur les pensées du sujet, non sur des actes dont la description contient une référence au corps, ce que semble n'avoir pas compris Gassendi.

Le fait que la tromperie du grand trompeur ait pour condition nécessaire la pensée et que la supposition d'une tromperie et la négation de la pensée rend la tromperie impossible, peut se traduire ainsi :

> « Quel que soit *x* et quel que soit *y*, le fait que *x* trompe *y* seulement si *y* pense, implique qu'il est certainement faux que *x* trompe *y* si l'on suppose à la fois que *x* trompe *y* et que *y* ne pense pas. »

L'énoncé qui précède est facilement traductible dans le langage du calcul des prédicats, et la syntaxe d'Imogen en offre une traduction presque limpide :

```
forall X. forall Y.((trompe(X,Y)=> pense(Y))=>
((trompe(X,Y) & ~ pense(Y))=> ~ trompe(X,Y)))
```

Vérifier qu'un tel énoncé est prouvable en logique intuitionniste devient dès lors un jeu d'enfant. Voici la copie du résultat du test de cette formule *via* Imogen :

```
joseph@joseph-Inspiron-530:~$ imogen prove
'forall X. forall Y.((trompe(X,Y)=> pense(Y))=>
((trompe(X,Y) & ~ pense(Y))=> ~ trompe(X,Y)))'
### Warning: using slow symbols ###
% SZS status Theorem for forall X. forall
Y.((trompe(X,Y)=>
pense(Y))=>((trompe(X,Y) & ~ pense(Y))=> ~
trompe(X,Y)))

The formula is true!
```

On peut aussi traduire la formule dans le langage usuel du calcul des prédicats du premier ordre et en donner la preuve à l'aide de la méthode exposée dans l'Appendice A de cet article :

***Théorème 2.1* —**

$$\vdash_i \ \forall x \forall y((Txy \to Py) \to ((Txy \wedge \neg Py) \to \neg Txy))$$

Démonstration. —
1. $\underline{?\forall x \forall y((Txy \to Py) \to ((Txy \wedge \neg Py) \to \neg Txy))}$
2. $\underline{?\forall y((Tay \to Py) \to ((Tay \wedge \neg Py) \to \neg Tay))}$ (1)
3. $\underline{?((Tab \to Pb) \to ((Tab \wedge \neg Pb) \to \neg Tab))}$ (2) ✔
4. $Tab \to Pb$ (3) ✔
5. $\underline{?((Tab \wedge \neg Pb) \to \neg Tab)}$ (3) ✔
6. ***Tab*** (5)
7. $\neg Pb$ (5)
8. $?\neg Tab$ (5)
9. ***?Pb*** (7)

10. ***?Tab*** (4) $\times$ 11. ***Pb*** (4) $\times$

□

On imagine aisément la réaction que peut provoquer la démonstration qui précède. Au « grand appareil » qu'est la fiction du grand trompeur, il semble que l'on substitue ici l'appareil de la formalisation logique avec ses subtilités pour lesquelles Descartes, comme le rappelle Guenancia [**6**, pp. 32-34], n'avait que mépris. Mais je crois qu'un tel rejet traduirait plus l'ignorance et l'incompréhension de la signification d'une telle preuve que les scrupules de l'historien. Ce que montre une telle preuve est au contraire précieux dans la perspective intuitionniste qui, incontestablement, est bien celle de Descartes. En effet, cette preuve démontre qu'il suffit d'accorder, d'une part, que « penser » est une condition nécessaire pour « être trompé » (ce qui est évidemment impliqué par la fiction du grand trompeur) et, d'autre part, que les constantes logiques de la conjonction, de l'implication et de l'absurde, ont une signification naturelle, universelle et évidente, pour que la premier résultat positif des *Méditations* soit, du point de vue intuitionniste, une *vérité logique* irréfutable. Il faut rappeler que le logicien intuitionniste est, en théorie de la démonstration, plus exigeant que le logicien classique car, comme je l'ai expliqué dans un autre article [**20**], toutes les vérités de la logique intuitionnistes sont des vérités analytiques (raison pour laquelle le tiers exclu n'est pas une vérité de la logique intuitionniste). Tel est bien le cas de la vérité du « Je pense » dans les *Méditations*. On se souvient que Gueroult [**7**, p. 62] a insisté sur le fait que la certitude *métaphysique* du *Cogito* est entendue par Descartes comme une certitude *scientifique*. La démonstration en logique intuitionniste du raisonnement de Descartes confirme ce point, en prouvant formellement ce qui peut se comprendre intuitivement : aucun état de la connaissance concevable n'est compatible avec l'hypothèse d'un individu qui se trompe, ou qui est trompé, mais qui ne pense pas.

Comme le souligne avec justesse Guenancia [**5**, p. 65], « il est impossible de douter de sa propre existence, de *chose qui pense* ». C'est donc bien la pensée et, avec la pensée en général, tout contenu de pensée au moment même de la représentation, qui est mis à l'abri du doute hyperbolique dans la seconde Méditation. Quelle que soit l'obscurité et la confusion d'une idée, le grand trompeur ne peut faire en sorte que je n'ai pas cette représentation au moment où je la pense. Il se peut que je ne me représente rien de réel, mais, de la même façon qu'il est impossible que je ne pense pas lorsque je doute, il est impossible que cette représentation ne soit rien, c'est-à-dire n'existe pas, quand je pense.

Mais on recherche désormais une idée qui puisse correspondre certainement à une chose existant hors du *Cogito*. Il est temps d'aborder ce qui est à mon avis la plus grande difficulté des

Méditations, celle de la preuve qui permet au *Cogito* de prouver que l'idée de Dieu qui est en lui à la fois *prouve l'existence* d'un être qui est hors de lui, et exprime correctement l'essence de Dieu. C'est l'objet de la troisième Méditation.

3. Méditation Troisième
Validité intuitionniste de la première preuve de l'existence de Dieu

La première preuve de l'existence de Dieu que Descartes donne dans les *Méditations* est dite « par les effets » parce qu'elle procède de l'idée que le *Cogito* a de Dieu pour prouver l'existence de Dieu qui est cause de cette idée. Dans la troisième Méditation, Descartes n'expose pas de manière abrupte cette preuve mais, de la même façon que pour la preuve qu'il donne du *Cogito*, s'efforce de préparer son lecteur à la compréhension de cette preuve. On ne fera pas ici l'analyse conceptuelle de cette preuve qui a longuement été développée par Gueroult [**7**, ch. V, pp. 154-247]. On s'efforcera au contraire d'en donner la traduction la plus concise et la plus précise possible afin de savoir, conformément à l'objectif poursuivi dans cet article, si cette preuve est exprimable et recevable en logique intuitionniste du premier ordre. La preuve de Descartes, indépendamment du travail conceptuel de préparation qui la précède, tient dans l'alinéa suivant :

> Partant il ne reste que la seule idée de Dieu, dans laquelle il faut considérer s'il y a quelque chose qui n'ait pu venir de moi-même. Par le nom de Dieu j'entends une substance infinie, éternelle, immuable, indépendante, toute connaissante, toute-puissante, et par laquelle moi-même, et toutes les autres choses qui sont (s'il est vrai qu'il y en ait qui existent) ont été créées et produites. Or ces avantages sont si grands et si éminents, que plus attentivement je les considère, et moins je me persuade que l'idée que j'en ai puisse tirer son origine de moi seul. Et par conséquent il faut nécessairement conclure de tout ce que j'ai dit auparavant, que Dieu existe. Car, encore que l'idée de la substance soit en moi, de cela même que je suis une substance, je n'aurais pas néanmoins l'idée d'une substance

> infinie, moi qui suis un être fini, si elle n'avait été mise en moi par quelque substance qui fût véritablement infinie.

Les six prémisses de cette preuve se répartissent en deux groupes de même taille. Le premier est constitué de trois axiomes (propositions 3.1 à 3.3) qui expriment les rapports évidents pour la pensée entre réalité des idées et causes des idées. Le second groupe est constitué de trois propositions (3.4 à 3.6) qui expriment des *faits* dont le *Cogito* ne peut douter. Toutes les propositions qui suivent jouent un rôle dans la preuve. Pour plus de clarté, elles sont aussitôt traduites par exprimées par une formule de la logique du premier ordre écrites dans la syntaxe d'Imogen. Il faut aussi préciser que la preuve qui suit fait usage de deux constantes d'individus qui n'appartiennent pas à la même catégorie grammaticale. En effet le terme *deo*, écrit en minuscules, fait référence au *concept* de Dieu, c'est-à-dire à l'idée innée que le sujet pensant conçoit. Ce sujet pensant est désigné par le terme de *cogito* écrit aussi en minuscules, pour se plier à la syntaxe d'Imogen. Pour que le sens de la preuve soit clair, il faut donc garder à l'esprit que *deo* désigne un *objet de pensée*, quand *cogito* désigne la représentation du *sujet pensant*. Ces précisions étant faites, on peut maintenant exprimer, formaliser et tester la preuve de Descartes.

Proposition 3.1. — *Si un être pensant a un concept quelconque, alors il existe quelque chose qui est la cause de ce concept*[4].

```
(forall Y.(a_un_concept(cogito,Y)=> exists X.
cause(X,Y)))
```

Proposition 3.2. — *Une chose ne peut être la cause d'une autre qu'à la condition que la cause ait au moins autant de perfections (ou de réalité) que l'effet*[5].

---

[4] La réalité objective d'une idée doit venir du Cogito ou d'une autre chose existante, car sinon cette réalité aurait pour cause le néant, ce qui est impossible.

[5] Descartes écrit : « c'est une chose manifeste par la lumière naturelle, qu'il doit y avoir pour le moins autant de réalité dans la cause efficiente et totale

```
(forall X. forall Y.(cause(X,Y)=>
a_au_moins_autant_de_perfections_que(X,Y)))
```

Proposition 3.3. — *Un être ne peut être la cause de ce qui est actuellement infini qu'à la condition d'être aussi actuellement infini.*

```
forall X. forall Y.((cause(X,Y) &
actuellement_infini(Y))=> actuellement_infini(X)))
```

Proposition 3.4. — *Le Cogito a un concept de Dieu.*

```
(a_un_concept(cogito,deo))
```

Proposition 3.5. — *L'infini actuel est une propriété du concept de Dieu.*

```
(actuellement_infini(deo))
```

Proposition 3.6. — *Il est absurde d'affirmer qu'il y a au moins autant de perfections dans le Cogito que dans le concept de Dieu*[6].

```
(~ a_au_moins_autant_de_perfections_que(cogito,deo))
```

que dans son effet : car d'où est-ce que l'effet peut tirer sa réalité sinon de sa cause ? »
On remarque que « perfection » est synonyme de « réalité » et que Descartes n'applique une relation d'ordre sur les représentations qu'en raison des relations de dépendances logiques ou grammaticales : la substance pouvant être conçue ou connue sans ses modes, elle est d'une perfection supérieure aux modes ; la représentation de Dieu qui est la représentation d'une substance actuellement infinie a plus de perfection que n'importe quelle autre représentation, et donc la cause de la représentation de Dieu doit avoir au moins autant de perfection, ou réalité, que cette représentation en contient. En conséquence aucune substance finie ne peut être la cause de la représentation de Dieu, ce qui exclut que le Cogito soit la cause de cette représentation.

[6] Puisque le sujet pensant doute, il est nécessairement imparfait et donc fini ; le concept de Dieu enveloppe en revanche une infinité de perfections.

La conclusion de la première preuve « par les effets » de l'existence de Dieu est la suivante :

Proposition 3.7. — *Ce n'est pas le Cogito qui est la cause du concept de Dieu, mais un être qui est actuellement infini.*

```
(~cause(cogito,deo) & exists
X.(actuellement_infini(X) & cause(X,deo)))
```

On parvient maintenant au théorème de cette section :

Théorème 3.8. — *La conjonction de chacune des expressions qui formalisent les propositions 3.1 à 3.6 implique, en logique intuitionniste du premier ordre, l'expression qui formalise la proposition 3.7 ; autrement dit* :

$$\vdash_i ((3.1) \wedge (3.2) \wedge (3.3 \wedge (3.4) \wedge (3.5) \wedge (3.6)) \to (3.7) \quad (11)$$

Du seul point de vue de la logique intuitionniste, la preuve « par les effets » que Descartes donne de l'existence de Dieu est donc irréprochable.

Test avec Imogen. — joseph@joseph-Inspiron-530:~$

```
imogen prove '((forall Y.(a_un_concept(cogito,Y)=>
> exists X. cause(X,Y))) &
> (forall X. forall Y.(cause(X,Y)=>
> a_au_moins_autant_de_perfections_que(X,Y))) &
> (forall X. forall Y.((cause(X,Y) &
actuellement_infini(Y))=>
> actuellement_infini(X))) &
> (a_un_concept(cogito,deo)) &
> (actuellement_infini(deo)) &
> (~
a_au_moins_autant_de_perfections_que(cogito,deo)))=>
> (~cause(cogito,deo) & exists
X.(actuellement_infini(X) & cause(X,deo)))'
### Warning: using slow symbols ###
% SZS status Theorem for ((forall
Y.(a_un_concept(cogito,Y)=>
exists X. cause(X,Y))) &
(forall X. forall Y.(cause(X,Y)=>
a_au_moins_autant_de_perfections_que(X,Y))) &
```

```
(forall X. forall Y.((cause(X,Y) &
actuellement_infini(Y))=>
actuellement_infini(X))) &
(a_un_concept(cogito,deo)) &
(actuellement_infini(deo)) &
(~
a_au_moins_autant_de_perfections_que(cogito,deo)))=>
(~cause(cogito,deo) & exists
X.(actuellement_infini(X) & cause(X,deo)))

The formula is true!
```

Test avec ileanCoP. — ileanCoP est, comme Imogen, un vérificateur automatique de théorèmes pour la logique intuitionniste du premier ordre[7], écrit en Prolog par Jens Otten. La syntaxe de ce programme est très proche de celle d'Imogen. Voici la copie du résultat sans surprise du test de (11) avec ileanCoP :

```
?- % nnf_mm_intu compiled 0.00 sec, 10,932 bytes
% /home/joseph/provers/prolog-
provers/ileancop_swi/ileancop_swi.pl
compiled 0.00 sec, 18,216 bytes
true.
?- prove(((all Y:(a_un_concept(cogito,Y)=>
ex X: cause(X,Y))) ,
(all X: all Y:(cause(X,Y)=>
a_au_moins_autant_de_perfections_que(X,Y))) ,
(all X: all Y:((cause(X,Y) ,
actuellement_infini(Y))=>
actuellement_infini(X))) ,
(a_un_concept(cogito,deo)) ,
(actuellement_infini(deo)) ,
(~
a_au_moins_autant_de_perfections_que(cogito,deo)))=>
(~ cause(cogito,deo) , ex X:(actuellement_infini(X)
, cause(X,deo)))).
| | | | | | | | | 0.009999999999999998
true
```

[7] http://www.leancop.de/ileancop/

De la même façon qu'une calculette donne le résultat d'une longue opération arithmétique, Imogen et ileanCoP permettent de savoir si une formule de la logique du premier ordre est ou n'est pas prouvable en logique intuitionniste. Aux esprits chagrins qui seraient tentés de répondre que la copie de ces résultats ne remplace pas une démonstration, on donne la possibilité de faire cette démonstration avec la méthode donnée dans l'Appendice A de cet article.

Pour aider encore un peu ceux qui souhaitent faire la démonstration de (11), en voici la traduction dans un formalisme très usuel :

$$\begin{aligned}((\forall y(Rcy \to \exists x Cxy))\land \\ (\forall x \forall y (Cxy \to Axy))\land \\ (\forall x \forall y((Cxy \land Iy) \to Ix))\land \\ Rcd\land \\ Id\land \\ \neg Acd) \to \\ (\neg Ccd \land \exists x(Ix \land Cxd))\end{aligned} \tag{9}$$

Le code LaTeX de cette formule est le suivant :

```
((\forall y(Rcy\to \exists x Cxy))\land(\forall x
\forall y (Cxy \to Axy))
\land (\forall x \forall y((Cxy \land Iy)\to Ix))
\land
Rcd \land Id \land \neg Acd)\to (\neg Ccd \land
\exists x(Ix \land Cxd))
```

Il suffit de copier ces trois lignes de code qui précèdent et de les soumettre au Tree Proof Generator, écrit et mis en ligne[8] par Wolfgang Schwarz, pour obtenir la preuve de cette formule *en logique classique*. Le logiciel de Wolfgang Schwarz donne un arbre de 15 branches à la 381ème étape de son calcul qui démontre la validité de (11). L'accord des résultats obtenus par Imogen et ileanCoP ne laisse aucun doute sur le fait que (11) est un théorème de la logique intuitionniste du premier ordre. On laisse au lecteur qui a du temps à

[8] http://www.umsu.de/logik/trees/

perdre le soin de le prouver « à la main », avec la méthode de l'Appendice A ou une autre méthode de son choix. La conclusion qui suit porte sur les leçons philosophiques des démonstrations qui viennent d'être faites.

Conclusion

Vuillemin [**22**] a montré que la *Géométrie* de Descartes est d'inspiration intuitionniste. C'est à partir de ce constat qu'il a jugé que les *Méditations métaphysiques* expriment une position philosophique en raison du privilège accordé à ce qu'il appelle les *jugements de méthode*. Mais aucune réponse n'avait été apportée jusqu'à présent à la question de savoir si les preuves fondamentales des *Méditations* sont conformes ou non à la logique intuitionniste telle qu'elle a été rigoureusement définie par Heyting. Cet article démontre que tel est bien le cas pour les deux premières preuves des *Méditations*, à savoir la preuve du *Cogito* et la première preuve par les effets de l'existence de Dieu. Cette dernière conduit logiquement à se débarrasser du solipsisme et il n'y a donc aucune raison de penser que *Les Méditations* ne sont pas une œuvre philosophique logiquement stable, c'est-à-dire intuitionniste d'un bout à l'autre de la chaîne des raisons.

Le fait que la première preuve par les effets soit une preuve recevable en logique intuitionniste est un résultat remarquable dans l'histoire des preuves de l'existence de Dieu. Avant Descartes, Anselme avait donné une preuve comparable, longuement et profondément analysée par Vuillemin [**23**]. Cependant, comme l'a déjà montré Weingartner [**26**, 10] la preuve du Proslogion comporte une étape illégitime d'un point de vue intuitionniste et fait de cette preuve une preuve qui n'est valide qu'en logique classique. Pour le montrer, on peut reprendre l'analyse que Vuillemin [**23**, p.21] donne de la preuve d'Anselme :

> La preuve d'Anselme se distingue à la fois de la preuve ontologique qui déduit l'existence de la perfection absolue et de la preuve cartésienne par l'idée du parfait en moi. Elle ressemble à cette seconde preuve en ce qu'elle fait appel à une *cogitatio* et, en ceci, elle ressemble à une preuve par les effets. Mais cette *cogitatio* est réputée impossible et c'est de

> cette impossibilité que l'existence divine devra être déduite. En ce sens, on peut dire que la preuve du *Proslogion* est une preuve par les effets exclus en ce qu'elle part de ce qu'il est impossible de poser en relation à l'être dont cette impossibilité démontrera l'existence.

En définissant Dieu comme « ce qui est tel que rien de plus grand ne peut être pensé », Anselme parvient à la conclusion que Dieu existe en raison même de la définition qu'il en donne. En effet « ce qui est tel que rien de plus grand ne peut être pensé ne peut exister seulement dans l'intelligence », car sinon on pourrait penser un être plus grand qui serait aussi dans la réalité et lui serait supérieur, ce qui contredirait la définition. Donc ce qui est tel que rien ne peut être pensé de plus grand *ne peut pas ne pas* être pensé comme étant aussi dans la réalité. D'où la conclusion d'Anselme [**2**, Proslogion, ch. II, p. 180] : « Il existe donc, sans aucun doute, quelque chose dont on ne peut rien concevoir de plus grand, et dans l'intelligence, et dans la réalité ». On remarque ici que le raisonnement d'Anselme effectue une dérivation licite en logique classique mais proscrite en logique intuitionniste : le passage de l'énoncé modal « *ne peut pas ne pas* être » à la conclusion « est nécessairement ». Or cette dérivation revient à supprimer la double négation au profit de l'affirmation, ce que refuse de façon générale la logique intuitionniste. Certes, la formule

$$\neg\Diamond\neg p \rightarrow \Box p \tag{12}$$

est un théorème de **S4**, mais **S4** est un système de logique modale classique. Ce que rejette la logique intuitionniste, c'est tout d'abord la formule classique et non modale

$$\neg\neg p \rightarrow p \tag{13}$$

dont la traduction modale McKinsey-Tarski est dans **S4**[9]

$$\Box(\Box\neg\Box\neg\Box p \rightarrow \Box p) \tag{14}$$

[9] Voir sur cette question Bell *et alii* [**3**, pp. 212-214].

Or (14) *n'est pas* un théorème de **S4** mais de **S5**, comme l'attestent les résultats des tests réalisé à l'aide de « The LogicsWorkbench »[10] :

```
joseph@joseph-Inspiron-530:~$ lwb
Starting the LWB, please wait...

LWB - The Logics Workbench 1.1 linux
type 'help;' for help

> load (s4);
s4 user
s4> provable(box(box ~ box ~ box p -> box p));
false
s4> load (s5);
s5 s4 user
s5> provable(box(box ~ box ~ box p -> box p));
true
```

Il est donc incontestablement fondé, du point de vue de la logique intuitionniste, de rejeter la preuve d'Anselme. Or tel n'est pas le cas de la première preuve de l'existence de Dieu que Descartes donne dans la troisième Méditation[11], logiquement recevable dans une logique plus faible que la logique classique, elle est philosophiquement plus forte parce qu'il est plus difficile de rejeter un argument qui repose sur un noyau logique partagé aussi par les logiciens classiques.

Ce résultat signifie-t-il que la preuve de Descartes est dès lors susceptible de forcer la foi du logicien intuitionniste ? Évidemment non et une analyse du fonctionnement de la preuve permet de voir où le bât blesse. On a vu que la preuve de Descartes repose sur trois axiomes et trois énoncés factuels admis comme indubitables. La suppression d'un seul énoncé de l'ensemble des prémisses rend impossible la dérivation de la conclusion. On peut convenir de ne pas discuter des propositions 3.1 à 3.3 qui sont des définitions ou des axiomes. En revanche, la conjonction des prémisses 3.4 et 3.5 pose problème, car l'on peut évidemment contester le fait que l'on conçoit clairement et distinctement l'infini divin qui est au-delà de toute

[10] http://www.lwb.unibe.ch/

[11] Contrairement à ce que j'avais cru et développé dans un brouillon fautif.

possibilité d'accroissement, c'est-à-dire, pour reprendre la définition d'Anselme, « l'être tel que rien de plus grand ne peut être pensé ». Je peux admettre que j'ai un concept de Dieu, mais refuser de reconnaître que je conçois l'infini absolu et actuel *via* un tel concept. Je peux aussi reconnaître que par définition le concept de Dieu enveloppe l'infini actuel, mais nier qu'il y ait en moi l'intuition d'un tel concept. Hobbes l'a parfaitement compris lorsqu'il écrit et répète dans les *Troisièmes Objections* que Descartes ne prouve pas que nous avons une idée de Dieu. Descartes peut bien répondre qu'il est manifeste que nous en avons une, il montre simplement que la prémisse de cette preuve outrepasse la position intuitionniste au sens strict : c'est *un acte de foi réaliste*, celui-là même que rejettent les intuitionnistes contemporains qui refusent précisément d'admettre dans les mathématiques tout élément théologique. Vuillemin a vu avec précision ce point lorsqu'il écrit [**21**, chap. 1] :

> le seul recours à la causalité et à la correspondance dans la preuve de l'existence de Dieu à partir de son idée en moi ne saurait produire un cercle qui ne figurerait pas déjà dans le contenu même de l'idée de Dieu.

L'intuitionniste authentique peut admettre aisément qu'un enchaînement de dérivations logiques dépasse l'intuition mais reste légitime tant que l'on peut vérifier pas à pas, quel que soit le temps que cela puisse prendre, que chaque dérivation est logiquement recevable. En cela les outils d'aide à la vérification des théorèmes sont les bienvenus, comme on a pu le voir ; mais la présence au sein de raisonnements logiquement corrects d'éléments absolument non constructifs, non décidables ou non vérifiés, ou encore non intuitifs, frappe aussitôt le résultat de la preuve intuitionniste du même doute que tout ce qui relève de preuves non constructives dans les mathématiques classiques. S'il est démontrable que la première preuve de l'existence de Dieu que donne Descartes est recevable du point de vue intuitionniste, il n'est ni prouvé ni attesté que nous ayons le concept de Dieu que Descartes prête au Cogito. Ce dernier point est condamné à rester une question de foi, et non de logique.

Appendice A
La logique intuitionniste du premier ordre

Le symbolisme utilisé dans cet article est standard, tant pour la constante du vrai $\top$, du faux $\bot$, de la négation $\neg$, de la conjonction $\wedge$, de la disjonction $\vee$, de l'implication $\rightarrow$, la quantification universelle $\forall$ et la quantification existentielle $\exists$. On a aussi adopté la convention relâchée mais commode qui consiste à adopter les dernières lettres de l'alphabet, la plupart du temps x, y z, comme variables d'individus et les premières lettres de l'alphabet a, b, c, d, comme symboles de constantes d'individus.

***Définition A.1.* — Négation intuitionniste**

$$\neg A \Leftrightarrow A \rightarrow \bot \qquad (Def.\neg)$$

***Définition A.2.* — Modèles de Kripke**

Dans l'ensemble des formules de la logique intuitionniste du premier ordre, un modèle de Kripke est une structure $K = (|K|, \leq, \rho, \Vdash)$ tel que, *intuitivement* :

– $|K|$ **modélise le temps**, avec un minimum ρ (un "maintenant"),
– $\leq$ une relation d'ordre partiel (préordre) pour définir **l'ordre du temps**, et
– $\Vdash$ une relation dite de "forcing" (réalisation) jouant le rôle de *fonction d'évaluation* d'une formule A à l'instant α.

"$\alpha \Vdash A\alpha$" dit "A est vraie à l'instant α".

On peut alors étendre la relation de forcing à toutes les formules :

(1) Si $\alpha \Vdash A$ et si $\alpha \leq \beta$, alors $\beta \Vdash A$(propriété de "monotonie" ou de "persistance" : le vérifié ou le prouvé reste toujours vérifié ou prouvé);
(2) $\alpha \Vdash \top$, $\alpha \nVdash \bot$;
(3) $\alpha \Vdash A \wedge B$ si et seulement si $\alpha \Vdash A$ **et** $\alpha \Vdash B$;

(4) $\alpha \Vdash A \vee B$ si et seulement si $\alpha \Vdash A$ **ou** $\alpha \Vdash B$;

(5) $\alpha \Vdash A \to B$ si et seulement si pour tout β tel que $\alpha \leq \beta$, $\beta \Vdash A$ implique $\beta \Vdash B$;

(6) $\alpha \Vdash \neg A$, si et seulement si quel que soit β tel que $\alpha \leq \beta$, $\beta \nVdash A$ (i.e. $\beta \Vdash A$ n'est pas le cas).

(7) **Remarque :** $\alpha \nVdash A$ *n'implique pas* $\alpha \Vdash \neg A$ mais dit seulement "A **n'est pas prouvée à l'instant** α (*mais peut-être plus tard !*)".

Définition A.3. — **Procédure de preuves**

La procédure de preuve utilisée est celle donné par Bell, DeVidi et Solomon [**3**, chap. 5, pp. 184-223] qui a l'immense mérite d'être actuellement la plus simple à utiliser.

– **Règles de transformation des formules** : données dans la Table 1 :

Table 1. Règles intuitionnistes pour les arbres de Beth en logique du premier ordre

	Disjonction	Conjonction	Implication
Affirmée	$A \vee B$ $A \quad B$	$A \wedge B$ A B	$A \to B$ $?A \quad B$
Non affirmée (douteuse)	$?(A \vee B)$ ✔ $?A$ $?B$	$?(A \wedge B)$ ✔ $?A \quad ?B$	$\underline{?(A \to B)}$✔ A $?B$

	Equivalence	Négation
Affirmée	$A \leftrightarrow B$ $A \quad ?A$ $B \quad ?B$	$\neg A$ $?A$
Non affirmée (douteuse)	$?(A \leftrightarrow B)$ ✔ $?(A \to B) \quad ?(B \to A)$	$\underline{?\neg A}$ ✔ A

Instanciation de $\forall$	Instanciation de $\exists$	Instanciation de $?\forall$	Instanciation de $?\exists$
$\forall x Fx$ Fc (c non inédite : pouvant être reprise dans la liste des lettres déjà utilisées)	$\exists x Fx$ ✔ Fc (c inédite : ne pouvant être dans la liste des lettres déjà utilisées.)	$\underline{?\forall x Fx}$ ✔ $?Fc$ (c inédite)	$?\exists x Fx$ $?Fc$ (c non inédite)

– **Règles de cochage** : On appose une coche ✔ pour indiquer que l'on désactive la formule une fois celle-ci transformée par l'application de la règle. Les formules qui ne sont jamais cochées sont les formules affirmées (celles qui ne sont pas précédées de ?) et qui sont *transportables* de haut en bas à travers les traits horizontaux qui marquent le passage d'un état de la connaissance à un autre : leur vérité persiste à travers ces états.
– **Règle de transport** Il est permis de transporter n'importe quelle formule qui n'est pas précédée du signe ? à travers les lignes horizontales introduites par les règles $? \rightarrow$, $?\neg$ et $?\forall$.
– **Règle de clôture** : Pour qu'une branche soit fermée il faut deux conditions :

(1) l'affirmation d'une formule quelconque *P* et sa non-affirmation ?*P* apparaissent dans la même branche,

(2) que *P* et ?*P* ne soient pas séparés par un trait horizontal.

– **Règle de la fourche** Bell *et alii* [**3**, p. 199] notent qu'il est important de remarquer que l'application répétée des règles $? \rightarrow$ et $?\neg$ dans une même branche apparaissant à un même état du savoir - donc dans une branche qui n'est pas coupée par un trait horizontal - doit nécessairement entraîner l'introduction de lignes horizontales et séparées pour chaque application des règles.

N.B. : Dans un tel cas il suffit qu'*une seule branche de la fourche* soit fermée pour que *la branche qui a donné naissance à la fourche* soit aussi fermée (c'est-à-dire contradictoire), comme le montre l'exemple A.6, où l'on doit considérer que c'est la branche initiale (1. à 4.), celle qui a donné naissance à la fourche, qui est complètement fermée parce qu'elle conduit à une situation contradictoire.

Exemple A.4. —

$$\nvdash_i (A \rightarrow B) \vee (B \rightarrow A)$$

Démonstration. —

$$
\begin{array}{c}
1.\ ?((A \rightarrow B) \vee (B \rightarrow A)) \\
2.\ ?(A \rightarrow B)\ _{(1)}\ ✔ \\
3.\ ?(B \rightarrow A)\ _{(1)}\ ✔
\end{array}
$$

$$
\begin{array}{cc}
4.\ \overline{A}\ _{(2)} & 6.\ \overline{B}\ _{(3)} \\
5.\ ?B\ _{(2)} & 7.\ ?A\ _{(3)}
\end{array}
$$

□

Remarque A.5. — Sans la règle de la fourche, la règle de transport permettrait la clôture de la branche et rendrait donc incorrect le test de la formule $(A \rightarrow B) \vee (B \rightarrow A)$.

***Exemple A.6.* —**

$$\vdash_i \neg A \rightarrow \neg(A \wedge \neg B) \vee \neg B \tag{15}$$

Démonstration. —

$$
\begin{array}{c}
1.\ \underline{?(\neg A \rightarrow (\neg(A \wedge \neg B) \vee \neg B)}\ ✔ \\
2.\ \neg A\ _{(1)} \\
3.\ ?\neg(A \wedge \neg B)\ _{(1)}\ ✔ \\
4.\ ?\neg B\ _{(1)}\ ✔
\end{array}
$$

$$
\begin{array}{cc}
5.\ \overline{A \wedge \neg B}\ _{(3)}\ ✔ & 6.\ \overline{B}\ _{(4)} \\
7.\ A\ _{(5)} & \\
8.\ \neg B\ _{(5)} & \\
9.\ \neg \boldsymbol{A}\ _{(2)} & \\
10.\ \boldsymbol{?A}\ _{(9)} & \\
\times &
\end{array}
$$

□

– Règle de lecture du résultat du test : Un arbre de réfutation intuitionniste est un test de validité d'une formule qui se lit quand et seulement quand l'arbre est totalement développé. Un arbre de réfutation est totalement développé quand tous les connecteurs $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$ ont été éliminés, quand toutes les formules actives ont été cochées, et quand la règle de transport n'est plus applicable. Si toutes les branches d'un arbre totalement développé sont fermées, alors la formule qui a fait l'objet du test est valide en logique intuitionniste ; sinon chaque branche ouverte donne un contre-modèle de Kripke de la formule et montre que la formule testée n'est pas prouvable en logique intuitionniste.

Bibliographie

[1] ALQUIÉ, F., 1987[1950]. *La découverte métaphysique de l'homme chez Descartes*, Paris : Presses Universitaires de France.

[2] ANSELME, (Saint), 1967. *Œuvres philosophiques*, Paris : Aubier, tr. fr. de Monologion-Proslogion-De Veritate-De Libero-Arbitrio-De Concordia-De Voluntate, par P. Rousseau.

[3] BELL, J.L., DEVIDI, D. & SOLOMON, G., 2001. *Logical Options : An Introduction to Classical and Alternative Logics*, Peterborough (Canada) : Broadview Press.

[4] BROUWER, L.E.J., 1908. « De Onbetrouwbaarheid der logische Principes », Tijdrschrift voorWijsbegeerte, p. 152-8, en. tr. in L.E.J. Brouwer *Collected Works*, I, 1975, pp. 107-11 ; trad. fr. in J. Largeault (ed.), 1912. *Intuitionnisme et théorie de la démonstration*, Paris : Vrin, pp. 15-23.

[5] GUENANCIA, P., 1996. *Descartes - Bien conduire sa raison*, Paris : Gallimard.

[6] —, 2000. *Lire Descartes*, Paris : Gallimard, Folio Essais.

[7] GUEROULT, M., 1968. Descartes selon l'ordre des raisons, vol. 1, Paris : Aubier -Montaigne.

[8] GUEROULT, M., 1968. Descartes selon l'ordre des raisons, vol. 2, Paris : Aubier -Montaigne.

[9] HINTIKKA, J. « Cogito, Ergo Sum: Inference or Performance ? », in Lambert, K. [**11**], p. 145-170.

[10] KIJANIA-PLACEK, K & WOLENSKI, J. (éd.), 1998. *The Lvov-Warsaw School and Contemporary Philosophy*, Synthese Library, vol. 273, Kluwer Academic Publishers,.

[11] LAMBERT, K. (éd.), 1991. *Philosophical Applications of Free Logic*, Oxford (NY) : Oxford University Press.

[12] — , 1997. Free Logics, Their Foundations, Character, and Some Applications Thereof, ProPhil – Projekte zur Philosophie, vol. 1, Sankt Augustin : Academia Verlag.

[13] MCLAUGHLIN, S. & PFENNING, F., 2009. « Efficient Intuitionistic Theorem Proving with the Polarized Inverse Method », in Cade (Schmidt, R.A., éd.), *Lecture Notes in Computer Science*, vol. 5663, Springer, p. 230–244.

[14] ONG-VAN-CUNG, K.S. (éd.), 1999. *Descartes et la question du sujet, Débats philosophiques*, Paris : Presses Universitaires de France.

[15] PARIENTE, J.-C. « La première personne et sa fonction dans le Cogito », in Ong-Van-Cung, K.S. [**14**] —, p. 11-48.

[16] SPINOZA, B., 1954. *Oeuvres complètes*, Paris : Gallimard, Bibliothèque de la Pléiade. Traduction, présentation et notes par Callois, Francès et Misrahi.

[17] STEINER, M. 1979. « Cartesian Scepticism and Epistemic Logic », *Analysis* **39**, no. 1, p. 38–41.

[18] VAN ATTEN, M. « The Development of Intuitionistic Logic », in *Stanford Encyclopedia of Philosophy, Stanford University*, `http://plato.stanford.edu/entries/intuitionistic-logic-development/`, 2008 revision 2009.

[19] VIDAL-ROSSET, J., 2005. « Intuitionnisme (logique, philosophie) », in Dictionnaire des Idées, *Encyclopædia Universalis*, p. 422-424.

[20] —, 2011. « L'argument de Russell-Tennant », in *Autour des Principia Mathematica*, B. Russell et A. N. Whitehead 1910-1913 (Dijon) (A. Guay, éd.), Éditions universitaires de Dijon, p. 149-177.

[21] VUILLEMIN, J. – « Etre et Choix - Eléments de philosophie réaliste », Manuscrit inédit des Archives Vuillemin, Université de Lorraine, Archives Poincaré, UMR 7117 du CNRS, Nancy.

[22] —, 1987 [1960]. *Mathématiques et métaphysique chez Descartes*, Paris : Presses Universitaires de France.

[23] —, 1971. *Le Dieu d'Anselme et les apparences de la raison*, Paris : Aubier Montaigne.

A non reductionist logicism with explicit definitions

Pierre Joray
Université de Rennes 1

Abstract

This paper introduces and examines the logicist construction of Peano Arithmetic that can be performed into Leśniewski's logical calculus of names called Ontology. Against neo-Fregeans, it is argued that a logicist program cannot be based on implicit definitions of the mathematical concepts. Using only explicit definitions, the construction to be presented here constitutes a real reduction of arithmetic to Leśniewski's logic with the addition of an axiom of infinity. I argue however that such a program is not reductionist, for it only provides what I will call a picture of arithmetic, that is to say a specific interpretation of arithmetic in which purely logical entities play the role of natural numbers. The reduction does not show that arithmetic is simply a part of logic. The process is not of ontological significance, for numbers are not shown to be logical entities. This neo-logicist program nevertheless shows the existence of a purely analytical route to the knowledge of arithmetical laws.

1. Logicism and reductionism

Since Frege, Whitehead and Russell, logicism has been widely conceived as a program for the reduction of arithmetic to pure logic. If the goal of the intended reduction is clear and can be summarized by the short claim that arithmetic is nothing but logic, it is nevertheless far from easy to describe accurately what such a reduction is and what is its very significance. Technically speaking, the reducibility of a theory S_1 to a theory S_2 lies on the possibility to prove in S_2 the axioms of S_1 by means of explicit definitions of the primitive terms of S_1 in the language of S_2. But this does not mean that S_1 *is nothing but* S_2. It only means that S_1 can be *interpreted* in S_2. However, in this technical sense, the significance of a reduction's existence is not inconsiderable: first, it gives a consistency proof of S_1 relative to the consistency of S_2; second, it shows that certain S_2-entities can play the role of the objects of S_1, even if it does not guarantee that these two groups of entities are simply identical.

As we know, the kind of reduction logicists usually had in mind is much stronger than this technical process of reduction. For them, the definition of "number" was supposed to grasp the very notion of number. For original logicists, the aim of the search of a reduction of arithmetic to pure logic was actually to provide the most basic mathematical theory with a reliable epistemic foundation. Our mathematical knowledge would have been strongly secured if numbers had been shown to be logical entities, with properties depending only on basic logical laws.

Nevertheless, this form of logicism was a failure: Frege was faced with Russell's antinomy and the authors of the *Principia Mathematica* were forced to enlarge their logical basis with three non-logical axioms. After this historical impasse, the only way for logicism was the search of the weakest addition to pure logic allowing the reduction of Peano-Dedekind Arithmetic (PA), while preserving the epistemic component of the original program.

We know today, first from C. Parsons (1965), but also from C. Wright (1983) and G. Boolos (1987), that Russell's paradox was not actually the death sentence of Frege's work on the foundation of mathematics. What is now called *Frege's Theorem* – the proof that the fundamental laws of arithmetic can be derived from second-order logic through the addition of a single proper axiom – is considered as an extremely interesting result for the philosophy of mathematics. This axiom, usually called *Hume's Principle*, is the formula which is discussed by Frege in his *Grundlagen*, just before the unfortunate introduction of extensions. It can be formulated as

$$\text{(HP)} \qquad (\forall F)(\forall G)(N(F) = N(G) \equiv F \infty G)$$

where $N(F)$ stands for "the cardinal number of the concept F" and ∞ for the relation of equinumerosity between concepts.

But C. Wright's and B. Hale's claim that Frege's Theorem is still a form of *logicism* is highly controversial (Hale & Wright 2001). Introducing "cardinal number" as a new primitive term, HP is clearly a proper axiom which cannot be said to be logical. According to Wright, though they are not purely logical truths, the laws of arithmetic are still shown to be analytic or purely conceptual by Frege's Theorem. The neo-Fregeans from St Andrews[1] indeed consider HP as an analytical truth. But it is then quite difficult to understand what they intend to mean when they argue that stipulating HP as true is meaning-constitutive of the expression $N(-)$. If it is (analytically) true, HP is a proposition and its truth does not depend at all on our stipulation. And if it is not a proposition, it is an open formula, with $N(-)$ as a free variable. Actually, the only way I can understand "I stipulate HP as true" is as "let me consider $N(-)$ with one of *these* meaning*s* (if any!) which satisfy the open formula in question".

HP is clearly considered as an *implicit definition* of $N(-)$. But the problems with implicit definitions abound. On the one hand, as any other additional axiom, they modify the whole system and can even lead to contradiction (like Frege's Basic Law V, which is exactly shaped like HP). But even if HP is consistent, it is too strong as a *definition* of a *single* proper term, for it also modifies the logical constants with which the term to be defined is explained. For example, due to its impredicative character, HP excludes all finite models. In other terms, it involves an axiom of infinity. From a purely proof-theoretical point of view, the addition of HP makes provable formulas which contains only logical terms and which are not theorems of pure logic. For example:

---

[1] See in particular: Wright (1983), Hale & Wright (2001), Ebert & Rossberg (2007).

$$(\exists x_1)(\exists x_2)(x_1 \neq x_2)$$

On the other hand, HP is too weak to be a *logicist* definition, for it does not warrant a definite and unique meaning for the concept to be introduced. The so-called "Caesar problem" is a consequence of this weakness. Saying, with neo-Fregeans, that open sentences of the form "$y = N(F)$" are only satisfied by objects falling under the identity conditions expressed by HP is not enough, for we cannot exclude that there is no (non-standard) "system of objects" (possibly including Julius Caesar) and satisfying HP however (even if Julius Caesar clearly does not fall under the identity conditions which follows from *our* intended interpretation of $N(-)$).

In spite of its elegance and very economical character, it is not clear whether Frege Arithmetic (FA = second-order logic + HP) is more *logicist* than Peano-Dedekind Arithmetic (PA). After all, Peano's axioms (as a whole) also constitute an implicit definition of "zero", "number" and "successor", explaining the related concepts in terms of pure logic. The unsolved Caesar problem shows that FA also invites the criticism Russell opposed to PA:

> We want our numbers to be such as can be used for counting common objects, and this requires that our numbers should have a *definite* meaning, not merely that they should have certain formal properties. (Russell 1919: 10).

But a reduction of a theory to another one does not signify that the reduced theory *itself* is derivable in (or a part of) the other one. It only means that the former is *interpretable* in the latter. Strictly speaking, the very mathematical notion of number cannot be defined in logic or in any other theory, for numbers would, then, have properties we are not ready to recognize as arithmetical ones. Mathematicians' positive integers cannot *be* such things as extensions or other logical objects abstracted from concepts; they cannot *be* classes of classes, or a certain kind of Zermelo-Fraenkel sets, neither – as suggested by P. Simons (2007) – properties of multitudes. In all these cases, numbers would have mathematically irrelevant properties, expressed by propositions which do not belong to arithmetic[2]. What I am ready to call "number" in mathematics is only one of these abstract and general entities which strictly satisfy all the theorems of arithmetic and no other.

In this perspective, what a logicist approach to a mathematical theory can provide is only what I will call a picture of this theory, that is to say a specific interpretation in which purely logical objects or constructions can play the role of mathematical notions or entities (numbers, in particular). Nevertheless, the existence of such a purely logical picture is far from pointless for the philosophy of mathematics. Naturally, it first gives a relative proof of consistency. But it especially provides an objective and conceptual path to arithmetical knowledge. A logicist picture gives such a secured epistemic justification, because it allows for a conceptual construction to replace the intuitive content and naïve notions which lead mathematicians, step by step, in the development of their practice, to the axiomatic characterization of their theory. The route thus constructed is epistemically secured, for the truth of the propositions it consists of depends only on logic.

The possibility of reinterpretations is today widely recognized by logicians and mathematicians as an essential advantage of axiomatic theories. Russell's above mentioned criticism was clearly overtaken by further developments of formal sciences. Nonetheless, his requirements – that our numbers can be used for counting common objects and that they have a definite meaning – are perfectly relevant relative to what I have called a picture of arithmetic. In order to provide the kind of justification I have just described, the picture must be materially adequate in Tarski's sense – it must present an adequate analysis of the naïve notion of number we use when counting concrete objects. On the other hand, it must also be definite in meaning. This requires the meaning of the defined terms to be fully determined by the meaning of logical constants. For this reason, the use of any implicit definition should be prohibited.

---

[2] As a basic example, if zero is defined as the empty set in ZF set theory, it receives the non-arithmetical property to be part of all the sets.

Frege's requirement that only logical constants occur in his Basic Laws is not followed by the neo-Fregeans. FA is undoubtedly a very nice theory to which arithmetic can be reduced. Nevertheless, it is neither arithmetic itself (as PA is), nor is it a good logicist picture of arithmetic, for it does not exclude reinterpretations of the proper term introduced by HP. For the latter condition to be satisfied, only explicit definitions must be used in the picture's construction.

In the following pages, I am going to show that a valuable logicist picture of arithmetic can be constructed from logic using only explicit definitions. This will be done without introducing extensions of concepts or classes – even as incomplete symbols or way of speaking – but on the ground of Stanisław Leśniewki's logical notion of name. Like plural terms do in natural languages, Leśniewki's names enable the expression of pluralities of things in the logical language.

2. A logic of names

When we assert a numerical statement like "there are five continents", according to Frege, we are speaking about a property of the concept *continent*. For Russell and Whitehead, it is to the class of continents that the property is asserted. Of course, neither the concept, nor the class can simply be said to be five. Before being analyzed by means of the logical relation of equinumerosity, the property in question can only be described as *having five objects falling under it* (for the concept), or *being a member of it* (for the class). On the other hand, *being five* is obviously not a property of the objects themselves: the continents *are* five, but none of them *is* five. According to P. Simons (2007), the property in question is a property of the "multitude" of continents (a notion he says to be akin to Husserl's "Vielheit" or Russell's "class as many"). But where is the expected solution? Like with "the concept of continent" or "the class of continents", "the multitude of continents" is obviously a *singular term*. The ordinary fact that we can use a single word or a single expression in order to refer to several objects seems to be mysterious for logicians as long as we do not postulate the existence of a single intermediate entity which has the (still mysterious?) virtue to gather together the things in question.

The idea underlying the logical picture of arithmetic to be presented hereafter is much more unsophisticated. Without trying to *explain* the one-to-many link between expressions and objects, one just *observes* that ordinary language involves expressions or words which are used to refer sometimes to a single thing (for example, "Cairo" or "the capital of Egypt"), sometimes to several things ("the African capitals", "horses") and also sometimes to nothing ("the capital of Africa", "Ulysses", "round circles"). The basic idea is to interpret numbers as *certain semantic properties of names*. Zero will then be depicted as the property of a name to be empty, one as the property of a name to be singular and three as the property of a name to refer to three things.

Leśniewski's calculus

Called *Ontology*, Leśniewski's system is grounded on such basic observations. It is an expansion of a quantified propositional calculus (called *Protothetic*). It has a single specific axiom introducing the constant copula *epsilon* and variables for names:

AxOnto:

$$\lfloor ab \rfloor \lceil a\varepsilon b \equiv (\lfloor \exists c \rfloor \lceil c\varepsilon a \rceil \wedge \lfloor cd \rfloor \lceil (c\varepsilon a \wedge d\varepsilon a) \supset c\varepsilon d \rceil \wedge \lfloor c \rfloor \lceil c\varepsilon a \supset c\varepsilon b \rceil) \rceil$$

The left-hand side of the equivalence '$a\varepsilon b$' is the general form of a singular proposition. It can be read as "*a* is *b*", or more precisely "the only object denoted by '*a*' is also denoted by '*b*' ". In other words, '$a\varepsilon b$' is truly asserted if and only if '*a*' stands for a singular name and '*b*' for a singular or plural name which denotes (possibly among others) the object denoted by '*a*'.

Definition rules

Among several peculiarities of Ontology, the system includes rules for stating explicit definitions of two kinds. Instead of stating definitions in the meta-language – like in the *Principia*, using the unspecified symbol '$=_{df}$' and introducing only convenient abbreviations – Leśniewski uses his primitive logical constants for expressing the equivalence relation between the *definiendum* (*Dum*) and the *definiens* (*Diens*)[3]. The first rule allows the introduction of propositional constants and functors while the second one allows the introduction of nominal constants or functors. The definitional equivalence is thus expressed by one of the two forms[4]:

$\lfloor v_1 \cdots v_n \rfloor \lceil Dum \equiv Diens \rceil$ Def$_S$ (propositional rule)

$\lfloor v_1 \cdots v_n \rfloor \lceil a\varepsilon Dum \equiv (a\varepsilon a \wedge Diens) \rceil$ Def$_N$ (nominal rule)

where 1. the left and right-hand sides of the equivalence involve the same (free) variables $v_1, \ldots, v_n$; 2. *Diens* is a formula with only primitive or already defined symbols; 3. *Dum* is of the following form, where # is the symbol to be defined and no symbol occur more than once:

Dum: $\#(v_1 \cdots v_i)(v_{i+1} \cdots v_j) \cdots (v_k \cdots v_n)$

As we will see, this general form of *Dum* relates to three possibilities. First, there can be no variable in *Dum*. The defined symbol is then either a constant proposition (with Def$_S$), or a constant name (with Def$_N$), like in the following examples:

D1. $\mathrm{T} \equiv \lfloor p \rfloor \lceil p \equiv p \rceil$ Def$_S$ (T : constant *true*)

D2. $\lfloor a \rfloor \lceil a\varepsilon \wedge \equiv (a\varepsilon a \wedge \sim (a\varepsilon a)) \rceil$ Def$_N$ ($\wedge$: empty name)

D3. $\lfloor a \rfloor \lceil a\varepsilon \vee \equiv (a\varepsilon a \wedge a\varepsilon a) \rceil$ Def$_N$ ($\vee$: universal name)

In the second case, the variables of *Dum* occur in a single pair of parentheses. The defined symbol is then a functor:

D4. $\lfloor ab \rfloor \lceil =\{ab\} \equiv (a\varepsilon b \wedge b\varepsilon a) \rceil$ Def$_S$

($=\{ab\}$: *a* is the same object as *b*)

D5. $\lfloor ab \rfloor \lceil \cong\{ab\} \equiv \lfloor c \rfloor \lceil c\varepsilon a \equiv c\varepsilon b \rceil \rceil$ Def$_S$

---

[3] About Leśniewskian internal definitions vs classical external and meta-linguistic definitions, see Joray (2005), (2006) & (2011).

[4] Where, in Leśniewski's notation, the first pair of square brackets expresses the universal quantifiers and the second pair contains the formula on which it is applied.

($\cong\{ab\}$: *a* and *b* have the same reference(s))

D6. $\lfloor a \rfloor \lceil 0\{a\} \equiv \sim \lfloor \exists b \rfloor \lceil b\varepsilon a \rceil \rceil$ Def$_S$ ($0\{a\}$: *a* is empty)

D7. $\lfloor a \rfloor \lceil 1\{a\} \equiv a\varepsilon a \rceil$ Def$_S$ ($1\{a\}$: *a* is singular)

D8. $\lfloor ab \rfloor \lceil \approx [\varphi\psi] \equiv \lfloor a \rfloor \lceil \varphi\{a\} \equiv \psi\{a\} \rceil \rceil$ Def$_S$

($\approx [\varphi\psi]$: φ and ψ are co-extensive)

D9. $\lfloor ab \rfloor \lceil a\varepsilon(b{\cdot}c) \equiv (a\varepsilon b \wedge a\varepsilon c) \rceil$ Def$_N$

(nominal intersection)

D10. $\lfloor ab \rfloor \lceil a\varepsilon(b+c) \equiv (a\varepsilon b \vee a\varepsilon c) \rceil$ Def$_N$ (nominal union)

D11. $\lfloor ab \rfloor \lceil a\varepsilon(b-c) \equiv (a\varepsilon b \wedge \sim (a\varepsilon c)) \rceil$ Def$_N$

(nominal complement)

In the last case, the variables of *Dum* split up into more than one pair of parentheses. The symbol to be defined is thus a multi-link or parametric functor, i.e. a functor-forming functor:

D12. $\lfloor ab \rfloor \lceil \cong\langle a\rangle\{b\} \equiv \cong\{ba\} \rceil$ Def$_S$

(parametric co-reference; $\cong\langle a\rangle$: denoting like *a*)

D13. $\lfloor ab \rfloor \lceil \varepsilon\langle a\rangle\{b\} \equiv b\varepsilon a \rceil$ Def$_S$

(parametric *epsilon*; $\varepsilon\langle a\rangle$: being one of the *a*'s)

Without going into the details, I will just underline certain aspects of the definition rules which are central for the understanding of the definition of numbers in the next section.

First, it has to be noticed that the definition rules allow introducing symbols of categories which are not previously available in the language. This is quite obvious in the case of multi-link or parametric functors. In D12, for example, the parametric functor of co-reference is defined on the basis of the usual identity binary relation. The parametric functor is the result of a different analysis of the same content: first, it applies to '*a*' and the result '$\diamond\cong\langle a\rangle$' expresses the nominal property "denoting-(exactly)-the-*a*'s"; this property can, then, be applied to a name '*b*', obtaining '$\cong\langle a\rangle\{b\}$' which expresses that '*b*' denotes (exactly) the *a*'s. This process of definition is very similar to a λ-abstraction, and, in Leśniewski's language, '$\cong\langle a\rangle$' is exactly what would be expressed by '$\lambda b.(\cong\{ba\})$' in λ-notation.

Secondly, Leśniewski's system is such that the definition of a functor in a new category allows the use of variables of that category and the binding of these variables by quantifiers.

The power of definition rules makes Ontology a strong analytical tool. Every semantic category can be reached progressively and the order of the formal language depends on its specific definitional development. One of Leśniewski's main achievements was his ability to elaborate completely explicit semantic and syntactic constraints in order to impose extensionality to every category and to avoid ambiguity and contradiction in the potentially infinite process of definition[5].

3. A logicist construction[6]

In the following construction, natural numbers are going to be depicted as *cardinal properties of finite names* (names which denote only a finite quantity of objects). Before going into the definition of the general notion of natural number, let me first consider how any particular natural number can be defined. Zero and one have already been introduced by definitions D6 and D7:

D6. $\lfloor a\rfloor\lceil 0\{a\}\equiv\sim\lfloor\exists b\rfloor\lceil b\varepsilon a\rceil\rceil$ $\quad$ Def_S

D7. $\lfloor a\rfloor\lceil 1\{a\}\equiv a\varepsilon a\rceil$ $\quad$ Def_S

In order to define the successor *n'* of a previously defined number *n*, one has to state that a name *a* has the number *n'* if and only if a name which denotes exactly the *a*'s excepted one of them has the number *n* :

[5] For a complete presentation of Leśniewski's Ontology, see Miéville (1984), (2004) and also the papers in Srzednicki & Rickey (1984).

[6] For the full presentation of the following logicist construction, with proofs and technical details, see Gessler, Joray, Degrange (2005: 73-137). The construction is partially inspired from Canty (1967).

D14. $\lfloor \varphi a \rfloor \lceil S(\varphi)\{a\} \equiv \lfloor b \rfloor \lceil b\varepsilon a \wedge \varphi\{a-b\} \rceil \rceil$ $\mathrm{Def_S}$ (successor)

From this, it is obvious that a symbol $\boldsymbol{n}$ for any natural number $n > 0$ can be introduced with a definition of the following form, where '$S(\cdots)$' is iterated *n* times:

$$\lfloor a \rfloor \lceil \boldsymbol{n}\{a\} \equiv S(S(\ldots S(0)\ldots))\{a\} \rceil \qquad \mathrm{Def_S}$$

Let me turn now to the definition of equinumerosity. Two names will be said to be equinumerous if and only if there is a one-to-one correspondence between the references of the first name and the references of the other one. In other terms, the one-to-one relation must have the first name as its *domain* and the second name as its *co-domain*. Let then consider the following preliminary definitions of *one-to-one relation*, *domain* and *co-domain* of a relation:

D15.
$\lfloor R \rfloor \lceil 1\text{-}1[R] \equiv \lfloor abc \rfloor \lceil (R\{ac\} \wedge R\{bc\}) \vee (R\{ca\} \wedge R\{cb\}) \supset =\{ab\} \rceil \rceil$
$\mathrm{Def_N}$

D16. $\lfloor aR \rfloor \lceil a\varepsilon Dom\langle R\rangle \equiv (a\varepsilon a \wedge \lfloor \exists b \rfloor \lceil R\{ab\} \rceil) \rceil$ $\mathrm{Def_N}$

D17. $\lfloor aR \rfloor \lceil a\varepsilon CoDom\langle R\rangle \equiv (a\varepsilon a \wedge \lfloor \exists b \rfloor \lceil R\{ba\} \rceil) \rceil$
$\mathrm{Def_N}$

And finally the definition of *equinumerosity*:

D18.
$\lfloor ab \rfloor \lceil \infty\{ab\} \equiv \lfloor \exists R \rfloor \lceil 1\text{-}1[R] \wedge \simeq\{Dom\langle R\rangle\, a\} \wedge \simeq\{CoDom\langle R\rangle\, b\} \rceil \rceil$
$\mathrm{Def_S}$

The cardinality of a name is, thus, simply defined as the parametric version of '∞':

D19. $\lfloor ab \rfloor \lceil \infty\langle a\rangle\{b\} \equiv \infty\{ba\} \rceil$ $\mathrm{Def_S}$

By the abstraction of '*b*' in '$\infty\{ba\}$', one gets the complex functor '$\infty\langle a\rangle$', which expresses the nominal property "denoting as many objects as *a*" or "having the cardinality of *a*". '$\propto\langle\cdots\rangle$' is then a parametric functor. When it is applied to a name, the result is a functor expressing the cardinal property of this name. As numbers are depicted as properties of names, it is natural to read '$\propto\langle a\rangle$' as "the cardinal number of *a*"[7] and the following theorem, which is easy to prove from D19, as the Leśniewskian version of HP:

$$\lfloor ab \rfloor\lceil \approx[\propto\langle a\rangle \propto\langle b\rangle] \equiv \propto\{ab\}\rceil$$

Contrary to the Fregean version of HP, here, the left-hand side does not express an identity between singular names, but an identity between nominal functors. Leśniewskian versions of Fregean "abstraction principles" are strictly predicative. This has important consequences on the construction. First, Leśniewskian versions of the "abstraction principles" never lead to contradiction. In particular, the analogue of Frege's Basic Law V is perfectly harmless and can be easily inferred from D12:

$$\lfloor ab \rfloor\lceil \approx[\simeq\langle a\rangle \simeq\langle b\rangle] \equiv \simeq\{ab\}\rceil$$

Second, the fact that abstraction's results are not designated as objects preserves the ontological neutrality of logic. The theorems of Leśniewski's calculus are logically true in the sense that they are true in all domains, including the empty domain. A consequence of this is that there will be no way to avoid the addition of an axiom of infinity for the derivation of all Peano's propositions.

From D19, the general definition of *cardinal number* can then be stated as:

D20. $\lfloor \varphi \rfloor\lceil Cn[\varphi] \equiv \lfloor \exists a \rfloor\lceil \approx[\propto\langle a\rangle\, \varphi]\rceil\rceil$ $\quad$ Def_S

In order to specify which cardinal numbers are natural numbers the definition of finite names is required. Like in Frege's *Grundlagen*, this will be done using the notion of inductivity: a name is said to be *inductive* if it has all the properties of the empty name that are preserved by the addition of a single denotation:

[7] This is of course only a way of speaking, for '$\propto\langle a\rangle$' is not the name of an object, but a symbol for a function.

D21.

$$\lfloor a \rfloor \lceil Ind\{a\} \equiv \lfloor \varphi \rfloor \lceil (\varphi\{\wedge\} \wedge \lfloor bc \rfloor \lceil (\varphi\{c\} \wedge 1\{b\}) \supset \varphi\{c+b\} \rceil) \supset \varphi\{a\} \rceil \rceil$$

$\mathrm{Def}_{\mathrm{N}}$

From this, *natural numbers* can be characterized as the cardinal numbers of finite names:

D22. $\lfloor \varphi \rfloor \lceil Nn[\varphi] \equiv (Cn[\varphi] \wedge \lfloor a \rfloor \lceil \varphi\{a\} \supset Ind\{a\} \rceil \rceil$ $\mathrm{Def}_{\mathrm{S}}$

D6, D14 and D22 are the respective definitions in Leśniewski's Ontology of Peano's primitive terms *zero*, *successor* and *number*. Peano's propositions I, IV and V are derivable from these definitions in pure Ontology

P_{I} $Nn[0]$ (zero is a number)

P_{IV} $\lfloor \varphi \rfloor \lceil Nn[\varphi] \supset \sim (\approx [S(\varphi)0]) \rceil$

(zero is not the successor of a number)

P_{V} $\lfloor P \rfloor \lceil (P[0] \wedge \lfloor \varphi \rfloor \lceil (Nn[\varphi] \wedge P[\varphi]) \supset P[S(\varphi)] \rceil) \supset \lfloor \psi \rfloor \lceil Nn[\psi] \supset P[\psi] \rceil \rceil$

(mathematical induction)

the remaining two propositions being only derivable in infinite Ontology:

P_{II} $\lfloor \varphi \rfloor \lceil Nn[\varphi] \supset Nn[S(\varphi)] \rceil$

(the successor of a number is a number)

P_{III} $\lfloor \varphi\psi \rfloor \lceil (Nn[\varphi] \wedge Nn[\psi]) \supset (\approx [S(\varphi)S(\psi)] \supset \approx [\varphi\psi]) \rceil$

(different numbers have different successors)

The proofs, which are quite long, can be found by the reader in Gessler, Joray, Degrange 2005: 75-137.

The full picture of Peano Arithmetic is thus constructed in a third-order development of infinite Ontology: a system of pure logic with the addition of an axiom of infinity.

One can notice that the dependence of Peano's propositions vis-à-vis the single non-logical axiom is not in the same dependence one can read about in the *Principia*. Here, not only P_{III}, but also P_{II} ("the successor of a number is a number") requires the existence of infinitely many objects. This is due to the fact that P_{II} cannot be read here as "ambiguous as to type", avoiding the artificial meaning of P_{II} in the *Principia*: *for every number* n*, there is a type* t *in which the successor of* n *(in fact the analogue of* n *for* t*) is a number*.

As it has been shown by Nadine Gessler[8], type ambiguity is not required to warrant the unity of all the higher-degree arithmetics that can be developed in Ontology. Anyhow, since what is to be constructed is not arithmetic *itself*, but a logical *picture* of it – an interpretation of general arithmetic in a system of certain definite logical entities – the classical question of the unity of the type hierarchy of arithmetics becomes almost superfluous.

Conclusion

Considering the above logicist construction of Peano Arithmetic as a logical picture of the mathematical theory, I claim that the existence of the reduction of PA to Leśniewski's Ontology does not inform us about the nature or the essence of numbers. The reduction does not show that arithmetic is a part of logic. The kind of foundation obtained through this process is not of ontological significance. But the possibility to reach arithmetical laws in the realm of logic has an epistemic value. Using only explicit definitions, it shows a purely analytic route to the knowledge of arithmetic.

Of course the need of an axiom expressing the existence of infinitely many objects can certainly be considered as a limitation of this program. But neither in common counting, nor in any application of arithmetic, does the assumption that there will always be enough available objects for the successor of a given number to be different from the number in question imply any commitment concerning the nature of the real world. The axiom of infinity is not an empirical hypothesis, but a conceptual assumption specifying the kind of idealization through which we apply arithmetic to specific concrete or abstract situations.

---

[8] Cf. Gessler, Joray, Degrange (2005: 9-36).

Hence, providing a purely conceptual content which guides us to Peano's axioms, the reduction to infinite Ontology gives an analytic justification for the adoption of these axioms as forming the basis for our coherent and applicable theory of pure mathematics.

References

BOOLOS, G., 1987. "The consistency of Frege's *Foundations of Arithmetic*", in Thomson J. (ed.). *On Being and Saying: Essays in Honor of Richard Cartwright*. MIT Press. 3-20. [Reprinted in Boolos 1998: 183-201].

BOOLOS, G., 1998. *Logic, Logic and Logic*. Cambridge (Mass.): Harvard Univ. Press.

CANTY, J. T., 1967. *Leśniewski's Ontology and Gödel Incompleteness Theorem*. PhD. Thesis. Univ. of Notre Dame.

EBERT, P. A., ROSSBERG M., 2007. "What is the Purpose of Neo-Logicism?", in Joray P. (ed.) *Contemporary Perspectives on Logicism and the Foundation of Mathematics*, *Travaux de Logique* **18**, Universités de Neuchâtel. 33-61.

FREGE, G., 1884. *Die Grundlagen der Arithmetik*. Breslau: Koebner.

FREGE, G., 1893. *Grundgesetze der Arithmetik*. Jena: Pohle Verlag.

GESSLER, N., JORAY, P., DEGRANGE, C., 2005. *Le logicisme catégoriel. Travaux de logique* **16**. Université de Neuchâtel.

HALE, B., WRIGHT, C., 2001. *The Reason's Proper Study. Essays towards a Neo-Fregean Philosophy of Mathematics*. Oxford: Clarendon.

JORAY, P., 2002. "Logicism in Leśniewski's Ontology", *Logica Trianguli* (Łódz, Nantes, Santiago de Compostella) **6**. 3-20.

JORAY, P., 2005. "Should Definitions be Internal?" In Bilkova M., Behounek L. (eds). *The Logica Yearbook 2004*. Praha: Filosofia. 189-199.

JORAY, P., 2006. "La définition dans les systèmes logiques de Łukasiewicz, Leśniewski et Tarski", in Pouivet R., Rebuschi M. (eds). *La philosophie en Pologne 1918-1939*. Paris: Vrin. 203-222.

JORAY, P., 2011. "Axiomatiques minimales et définitions : la thèse de Tarski sur le calcul biconditionnel", *Travaux de Logique* **20**, Universités de Neuchâtel et de Rennes 1. 57-83.

LEŚNIEWSKI, S., 1992. *Collected Works* (2 vol.). Surma S. J., Srzednicki J. T., Barnett D. I. (eds). Warszawa: PWN / Dordrecht: Kluwer.

MIEVILLE, D., 1984. *Un développement des systèmes logiques de Stanisław Leśniewski. Protothétique, Ontologie, Méréologie*. Berne: Peter Lang.

MIEVILLE, D., 2001-2004. *Introduction à l'œuvre logique de S. Leśniewski. I. La Protothétique, II. L'Ontologie*. *Travaux de Logique*. Hors série, Université de Neuchâtel

RUSSELL, B., 1919. *Introduction to Mathematical Philosophy*. London: Allen & Unwin. [quoted in 1971 ed. New York: Simon and Shuster].

SIMONS, P. M., 2007. "What Numbers Really Are", in Auxier, R. E. & Lewis, E. H. (eds). *The Philosophy of Michael Dummett*. La Salle: Open Court.

SRZEDNICKI, J. T. J., RICKEY, V. F. (eds), 1984. *Leśniewski's Systems: Ontology and Mereology*. Boston, The Hague: Nijhoff / Wrocław: Ossolineum.

WHITEHEAD, A. N., RUSSELL, B., 1927. *Principia Mathematica*. 2nd ed. Cambridge Univ. Press. [1st ed. 1910].

WRIGHT, C., 1983. *Frege's Conception of Numbers as Objects*. Aberdeen Univ. Press.

Ontologie et théorie des ensembles

François Lepage
Université de Montréal

Résumé

L'ontologie de Leśniewski est un calcul général des noms. Elle fut créée par Leśniewski pour apporter une solution naturelle au paradoxe de Russell en théorie naïve des ensembles. L'ontologie a été perçue par ses défenseurs et par ses adversaires comme une théorie incompatible avec la théorie des ensembles. Dans le présent texte, nous montrons que l'ontologie de Leśniewski permet, au contraire, de définir une théorie des ensembles qui coïncide avec la théorie de Zermelo-Fraenkel.

Il est bien connu que les systèmes de l'ontologie de Leśniewski sont des systèmes ouverts au sens où l'on peut toujours y introduire de nouveaux noms par des définitions thèses. Les noms sont de la catégorie N et sont de trois espèces : les noms propres qui dénotent des *individus*, les noms qui dénotent des *multiplicités* et les noms qui ne dénotent pas. Des exemples des premiers seraient Socrate, la Lune ou le plus petit nombre de la forme $2^n - 1$ qui n'est pas premier. Des exemples des seconds seraient les chevaux, l'actuel gouvernement canadien ou les étoiles de l'univers. Enfin, des exemples des troisièmes seraient Macbeth, le Père Noël ou le plus grand nombre premier.

On peut supposer ou non que le système possède des *primitifs* des trois espèces et qu'il a la capacité d'en créer de nouveaux. Par exemple, si P dénote les nombres premiers, H défini par $(\forall x)(x \,\varepsilon H \equiv (\exists y)(y \,\varepsilon P \wedge x = y + 1))$ dénotera les successeurs des nombres premiers.

Dans l'optique de redéfinir les ensembles à l'intérieur de l'ontologie, nous l'enrichissons d'un nouveau foncteur subtilement noté « {} » qui à partir de noms des trois espèces forme un nouveau nom de la première espèce (on suppose que « {} » est un nouveau symbole, qu'il ne faisait pas partie de l'alphabet de l'ontologie avant son introduction). C'est un créateur de noms propres que j'appelle l'*ensembleur*. Les propriétés spécifiques seront présentées plus loin. Pour l'instant, il s'agit simplement de créer de nouveaux noms propres qui possèdent toutes les propriétés communes à tous les noms propres. Tous les nouveaux noms ainsi obtenus le sont en utilisant le foncteur « {} » au plus un nombre *fini* de fois (même si ce nombre n'est pas borné) et, bien sûr, les définitions thèses. À l'aide de définitions appropriées, nous imposons des propriétés à ces nouveaux objets. Nous étudions ensuite les propriétés de ces ensembles *ontologiques*, nous intéressant particulièrement aux axiomes de Zermelo-Fraenkel.

1. L'essentiel de l'ontologie

La comparaison entre les systèmes de l'ontologie de Leśniewski et les théories des ensembles de la tradition Frege-Russell a donné lieu à de nombreuses publications. La question de la traduction des uns dans les autres n'a cependant pas, à ma connaissance, donné de résultats entièrement satisfaisants. Deux grandes voies s'offrent à nous. La première consiste à trouver un système axiomatique dans le cadre de la logique classique et à montrer que le système ainsi obtenu est la traduction des systèmes de l'ontologie. Cela pose certains problèmes car les systèmes classiques sont clos alors que les systèmes ontologiques sont ouverts. Cette difficulté ne doit cependant pas être insurmontable.

La seconde voie consiste à essayer de reconstruire la (une) théorie des ensembles dans le cadre de l'ontologie. C'est ce que je me propose de faire.

Je vais commencer par présenter minimalement l'ontologie en ne m'attardant que sur ce qui est essentiel pour comprendre la suite. J'utiliserai la notation classique pour les connecteurs logiques.

L'ontologie est une extension de la protothétique qui elle-même est un calcul propositionnel généralisé où l'on peut quantifier sur toutes les variables de toutes les catégories.

Définition des catégories de la protothétique

(i) S est une catégorie syntaxique ;
(ii) Si $X, X_1, ..., X_n$ sont des catégories syntaxiques, alors $X/X_1, ..., X_n$ est une catégorie syntaxique.

Les expressions de la protothétique sont celles qui possèdent une des propriétés suivantes :

(i) A est une variable de la catégorie S ;
(ii) $(A \equiv B)$ où A et B sont des expressions de la protothétique ;
(iii) $(\forall x_1)...(\forall x_n)A$ où A est une expression de la protothétique ;
(iv) toutes les expressions de la forme $N(v_1...v_n)$ où N est introduite par une définition de la forme $(\forall v_1)...(\forall v_n)(N(v_1...v_n) \equiv A(v_1, ..., v_n))$.

Exemple : définition de la conjonction

$(\forall p)(\forall q)(\wedge(pq) \equiv (\forall f)(p \equiv ((\forall r)(p \equiv f(r)) \equiv (\forall r)(q \equiv f(r)))))$
où f est une variable de la catégorie S/S.

Étant donné le caractère peu intuitif des expressions de la protothétique, je vous épargne la présentation du système axiomatique. Sachez seulement qu'il est fiable et complet.

L'ontologie est une extension de la protothétique. Outre la catégorie S, elle comporte la catégorie N, celle des noms.

Définition des catégories de l'ontologie

(i) S et N sont des catégories syntaxiques ;
(ii) Si $X, X_1, ..., X_n$ sont des catégories syntaxiques, alors $X/X_1, ..., X_n$ est une catégorie syntaxique.

L'ontologie contient également un nouveau foncteur primitif « ε » (à ne pas confondre avec « $\in$ » que nous allons définir plus loin).

« $a \varepsilon b$ » se lit « a est parmi les b » où les b sont soit plusieurs, un seul ou ne sont pas. ε est de la catégorie S/NN(« $a \varepsilon b$ » tient lieu de « $\varepsilon(ab)$ ».

Au système axiomatique de la protothétique on ajoute un axiome pour obtenir l'ontologie :

$$(\forall a)(\forall b)((a \varepsilon b) \equiv ((\neg(\forall c)\neg(c \varepsilon a) \wedge (\forall d)(\forall c)(((d \varepsilon a) \wedge (c \varepsilon a)) \supset (d \varepsilon c)) \wedge (\forall d)((d \varepsilon a) \supset (d \varepsilon b))))$$

Cela signifie exactement (Simons 1981) que (i) a n'est pas un nom vide, (ii) chaque a est un b et (iii) a est un individu.

Une conséquence immédiate est que $(a \varepsilon b)$ n'est jamais vrai si a est un nom vide ou a est le nom d'une multiplicité.

En utilisant les définitions thèses, on peut introduire les foncteurs suivants :

Quelque chose est un a

$$(\forall a)(ex(a) \equiv (\exists b)(b \varepsilon a))$$

Au plus une chose est un a

$$(\forall a)(sol(a) \equiv (\forall c)(\forall d)((c \varepsilon a) \wedge (d \varepsilon a) \supset (c \varepsilon d)))$$

Il y a une et une seule chose qui soit un a

$$(\forall a)(ob(a) \equiv (ex(a) \wedge sol(a)))$$

Toute chose qui est un a *est un* b

$$(\forall a)(\forall b)(a \subset b \equiv (\forall c)(c \varepsilon a \supset c \varepsilon b))$$

Utilisant ces notations et l'axiome de l'ontologie, on peut démontrer le théorème suivant qui exprime de façon claire ce que dit l'axiome :

$$(\forall a)(\forall b)((a \varepsilon b) \equiv ((ex(a) \wedge a \subset b \wedge sol(a)))$$

Avant de parler des ensembles, il vaut la peine de dire quelques mots sur la définition des classes dans l'ontologie. Il ne faut pas oublier que l'ontologie était, selon son inventeur, censée apporter une solution naturelle à la contradiction de Russell de la classe des classes qui ne s'appartiennent pas.

Sobosinski (1984) prête à Leśniewski deux notions de classes : la classe collective $Kl(a)$ et la classe distributive $el(a)$, les deux notions étant reliées par l'équivalence suivante :

$$(\forall a)(\forall b)(b \mathbin{\varepsilon} el(a) \equiv (\exists c)(a \mathbin{\varepsilon} Kl(c) \wedge b \mathbin{\varepsilon} c))$$

Le statut de cette équivalence n'est pas clair : elle est difficile à interpréter sans faire appel à l'intuition qu'elle est censée rendre explicite. Toujours est-il que si c est une expression à dénotation multiple, $Kl(c)$ est un nouvel individu et $el(Kl(c))$ est identique à c. Si a est un individu $Kl(a)$ est a et $el(Kl(a))$ est $el(a)$ qui est a.

2. Théorie des ensembles ontologiques

Oublions pour l'instant la théorie des classes de l'ontologie (nous y reviendrons plus loin) et penchons-nous sur la question de la définissabilité d'une théorie relativement standard des ensembles en apportant des modifications les plus mineures possible à l'ontologie.

La stratégie est simple. On introduit un nouveau foncteur, $\{\}$, l'ensembleur, régi par la *directive ensembliste* suivante. Pour tout *a* de la catégorie N, $\{\}a$ est un nouvel *individu* (on suppose que $\{\}$ est un nouveau symbole). À chaque étape de son développement, un système ontologique est ainsi enrichi d'un nombre indéfini de nouveaux noms obtenus non seulement par l'application de $\{\}$ mais par les définitions thèses utilisant entre autres les nouveaux noms obtenus par l'application de $\{\}$. Nous appelons ce nouveau système l'*ontologie élargie*.

Par exemple, si nous n'avons que deux noms propres primitifs dans notre système, soit *a* et *b*, nous pouvons non seulement définir c

$$(\forall x)(x \mathbin{\varepsilon} c \equiv (x \mathbin{\varepsilon} a \vee x \mathbin{\varepsilon} b))$$

mais également

$$(\forall x)(x \mathbin{\varepsilon} d \equiv (x \mathbin{\varepsilon} a \vee x \mathbin{\varepsilon} \{\}\, b))$$

ou encore

$$(\forall x)(x\,\varepsilon\, e \equiv (x\,\varepsilon\, a \vee x\,\varepsilon\,\{\}a \vee x\,\varepsilon\,\{\}\{\}c \vee x\,\varepsilon\,\{\}d))$$

où c est comme ci-dessus. La seule contrainte pour le moment est que pour tout x de la catégorie N, $\{\}x$ est également de la catégorie N et que $\{\}x$ est un individu.

Plus formellement, réservons l'appellation N aux noms *primitifs* de l'ontologie, c'est-à-dire ceux qui ne contiennent pas $\{\}$ et appelonsN' les noms de l'ontologie élargie ceux à qui sont tels que x est soit un terme primitif, soit un terme de la forme $\{\}x$ où x est un nom de l'ontologie élargie (encore une fois, x peut avoir été introduit par des définitions thèses de l'ontologie élargie).

L'égalité se définit de la manière suivante en ontologie :

$$(\forall x)(\forall y)(\forall z)((x \circ y) \equiv (z\,\varepsilon\, x \equiv z\,\varepsilon\, y))$$

Cette égalité est à distinguer de l'identité entre individus qui est l'égalité forte

$$(\forall x)(\forall y)((x = y) \equiv (x\,\varepsilon\, y \wedge y\,\varepsilon\, x))$$

Nous aurons besoin du nom vide.

La définition standard du nom vide en ontologie est la suivante :

$$(\forall x)(\forall y)(x\,\varepsilon\,\Lambda \equiv (x\,\varepsilon\, x \wedge \neg(x\,\varepsilon\, x)))$$

Il est clair que $a\,\varepsilon\,\Lambda$ est toujours faux.

Nous aurons besoin des deux axiomes suivants. On suppose que la quantification porte sur les noms et non sur les expressions d'ordre supérieur car on peut exprimer qu'une expression est de la catégorie N' en conditionnalisant :

$(\forall x)A(x)$ devient $(\forall x)(((x \circ \Lambda) \vee (\exists y)(y\,\varepsilon\, x)) \supset A(x))$

car $(x \circ \Lambda) \vee (\exists y)(y\,\varepsilon\, x)$. n'est vrai que si x est de la catégorie N'. Enfin, de manière abusive mais non problématique, nous dirons que A est l'extension de $\{\}A$.

E1 $(\forall A)((\{\}A = \{\}B) \equiv (A \circ B))$
E2 $(\forall A)(\{\}A \,\varepsilon\, \{\}A)$

E1 signifie que deux ensembles sont égaux ssi ils sont coextensifs.

E2 signifie que si A est un nom alors il y a un ensemble qui a A comme extension et cet ensemble est un individu. Malgré sa simplicité, E2 est relativement puissant : la collection des noms est fermée pour les ensembles. Chaque fois qu'un nom est introduit, un nouvel individu est introduit, soit l'ensemble ayant ce nom pour extension. Cela implique une définition inductive des ensembles qui n'est pas dans l'esprit Lesniewskien en autant que je puisse en juger. Par exemple, si a est de la catégorie N alors $\overbrace{\{\}...\{\}}^{n \text{ fois}}a$ est défini et est également un individu pour tout n fini.

Proposition

$a \,\varepsilon\, \{\}a$ est toujours faux.

Preuve

Si a n'est pas un individu, $a \,\varepsilon\, \{\}a$ est faux. Si a est un individu, comme $\{\}a$ est aussi un individu, par l'axiome de l'ontologie, $a \,\varepsilon\, \{\}a$ et $a\{\} \,\varepsilon\, a$ et donc $a = \{\}a$ ce qui est impossible : on aurait alors que $a = \overbrace{...\{\}\{\}}^{\text{un nombre infimi de fois}}a$ qui n'est pas une expression de l'ontologie.

Proposition

Aucune expression de la catégorie N ne contient plus qu'un nombre fini d'emboîtements de « $\{\}$ ».

Preuve

La construction de cette expression nécessiterait l'application du foncteur $\{\}$ un nombre infini de fois.

Appartenance

Définition de $\in$

$$(\forall x)(\forall y)(x \in \{\}y \equiv x \,\varepsilon\, y)$$

Si z n'est pas de la forme $\{\}w$, $x \in z$ est faux.

Cette définition appelle quelques commentaires. Elle est due à Peter Simons (1981, 189) bien qu'on la retrouve de façon implicite chez Küng et Canty (1970, 178). Reprenons notre exemple du e ci-dessus. On a bien $x \in \{\}e \equiv x \,\varepsilon\, e$ est vrai ssi $x \in \{\}e \equiv (x \,\varepsilon\, a \vee x \,\varepsilon\, \{\}a \vee x \,\varepsilon\, \{\}\{\}c \vee x \,\varepsilon\, \{\}d)$ ce qui correspond bien à l'intuition « ontologique ».

Proposition

$x \in x$ est toujours faux

Preuve

Pour que cette proposition soit vraie, il faut que $x = \{\}y$ pour un certain y. On a alors $\{\}y \in \{\}y$. Cela entraine que $\{\}y \,\varepsilon\, y$. Si un des y est $\{\}y$, alors le nombre fini n d'emboîtements de $\{\}$ de y est supérieur au nombre $m \geq n+1$ d'emboîtements de $\{\}$ de $\{\}y$ ce qui est impossible.

Définition de l'ensemble vide $\varnothing$

$\varnothing$ est une abréviation pour $\{\}\Lambda$

Proposition

Pour tout a, $a \in \varnothing$ est faux.

Preuve

$a \in \varnothing$ ssi $a \in \{\}\Lambda$ ssi $a \,\varepsilon\, \Lambda$ ssi $a \,\varepsilon\, a$ et $\neg(a \,\varepsilon\, a)$ ce qui est toujours faux.

3. Quelques propriétés des ensembles ontologiques

Nous allons montrer que les axiomes élémentaires de Zermelo-Fraenkel, c'est-à-dire les axiomes autres que l'axiome de l'infini et le schéma d'axiome de remplacement sont valides. Dans ce qui suit, les lettres majuscules dénotent des ensembles c'est-à-dire sont de la catégorie N et sont de la forme $\{\}a$ pour un certain a de la catégorie N.

Axiome d'extensionalité

$$(\forall x)(x \in A \equiv x \in B) \equiv (A = B)$$

Comme A et B sont des ensembles, A est $\{\}w$ et B est $\{\}z$ pour quelques w et z. On a alors $(\forall x)(x \in A \equiv x \in B)$ est vrai ssi $(\forall x)(x \,\varepsilon\, w \equiv x \,\varepsilon\, z)$ est vrai ssi $w \circ z$ ssi $\{\}w = \{\}z$ ssi $A = B$.

Axiome de séparation

$$(\exists B)(\forall x)(x \in B \equiv x \in A \wedge x \,\varepsilon\, z)$$

Ici, nous avons choisi de représenter le $\varphi(x)$ traditionnel par $x \,\varepsilon\, z$. Cela est justifié par la règle générale de définition ontologique de type ontologique (Miéville 2004, 64). Cet axiome est trivialement vrai parce que si A est un ensemble, A est $\{\}w$ pour quelque w et est tel que si $x \,\varepsilon y \equiv (x \,\varepsilon\, w \wedge x \,\varepsilon\, z)$ alors $(\forall x)(x \in \{\}y \equiv x \in A \wedge x \,\varepsilon\, z)$.

Axiome de la paire

$$(\exists A)(\forall x)(x \in A \equiv (x \in y \vee x \in z))$$

Supposons que A soit $\{\}w$.

$$(\exists A)(\forall x)(x \,\varepsilon\, w \equiv (x \in y \vee x \in z))$$

Si x et y ne sont pas des ensembles, w est un nom vide et A est $\varnothing$.
Si x ou y sont des ensembles, A existe par la règle générale de définition ontologique de type ontologique.

Axiome de la somme

$$(\exists C)(\forall x)(x \in C \equiv (\exists B)(x \in B \wedge B \in A))$$

Supposons que A soit $\{\}w$. Soit v tel que

$$(\forall x)(x \,\varepsilon\, v \equiv (\exists u)(x \,\varepsilon\, u \wedge \{\}u \,\varepsilon\, w)).$$

On vérifie facilement que

$$(\forall x)(x \in \{\}v \equiv (\exists B)(x \in B \wedge B \in A))$$

Axiome de l'ensemble puissance

$$(\exists B)(\forall C)(C \in B \equiv C \subseteq A)$$

$C \subseteq A$ est une abréviation pour $(\forall x)(x \in C \supset x \in A)$
Supposons que A soit $\{\}w$. On définit v de la façon suivante :

$$(\forall u)(\{\}u \,\varepsilon\, v \equiv ((\forall x)(x \,\varepsilon\, u \supset x \,\varepsilon\, w)))$$

$\{\}v$ est tel que $(\forall C)(C \in \{\}v \equiv C \subseteq A)$

Axiome de régularité

Cet axiome s'énonce comme suit :

$$A \neq \varnothing \supset (\exists x)(x \in A \wedge (\forall y)(y \in x \supset y \notin A)).$$

Supposons qu'il y ait un A qui rende cet axiome faux et donc un A non vide qui rend vrai $(\forall x)(x \in A \supset (\exists y)(y \in x \wedge y \in A))$.

Comme A est un ensemble non vide soit $a \in A$. $a \in A \supset (\exists y)(y \in a \wedge y \in A)$ et $a \in A$ entraîne $(\exists y)(y \in a \wedge y \in A)$. Mais $(\exists y)(y \in a \wedge y \in A)$ entraîne qu'il y a un c_1 tel que $c_1 \in a$ et $c_1 \in A$.

$c_1 \in A \supset (\exists y)(y \in c_1 \wedge y \in A)$ et $c_1 \in A$ entraîne $(\exists y)(y \in c_1 \wedge y \in A)$. Mais $(\exists y)(y \in c_1 \wedge y \in A)$ entraîne qu'il y a un c_2 tel que $c_2 \in c_1$ et $c_2 \in A$.

.

.

.

$c_n \in A \supset (\exists y)(y \in c_n \wedge y \in A)$ et $c_n \in A$ entraîne $(\exists y)(y \in c_n \wedge y \in A)$. Mais $(\exists y)(y \in c_n \wedge y \in A)$ entraîne qu'il y a un c_{n+1} tel que $c_{n+1} \in c_n$ et $c_{n+1} \in A$.

.

.

.

Donc si A falsifie l'axiome de régularité, il existe une chaine infinie $a \ni c_1 \ni c_2 \ni c_3$...

Mais d'après la définition de $\in$, si $x \in y$, y possède une paire de « $\{\}$ » de plus que le z tel que $z \varepsilon x$ qui en possède le plus. Donc une telle chaine infinie n'existe pas.

Schéma d'axiomes de remplacement

Le schéma d'axiomes de remplacement est la suite dénombrable d'énoncés (Krivine 1969) :

$$(\forall x_1)...(\forall x_n)(((\forall x)(\forall y)(\forall y')((A(x,y,x_1,...,x_n) \wedge A(x,y',x_1,...,x_n)) \supset (y = y'))) \supset (\forall t)(\exists w)(\forall v)(v \in w \equiv (\exists u)(u \in t \wedge A(u,v,x_1,...,x_n))))$$

La forme logique de cet énoncé est : si $A(x,y,x_1,...,x_n)$ est fonctionnelle en x y, alors l'image d'un ensemble par cette relation fonctionnelle est un ensemble. Le problème ici est de définir une relation fonctionnelle en général. Si on se limite aux relations fonctionnelles élémentaires que l'on peut définir à partir de la notion de paire, ce schéma est trivialement valide :

$$(\forall x)(\forall y)(\forall z)(\exists w)(y \,\varepsilon\, w \equiv y \,\varepsilon\, y \wedge x \,\varepsilon\, z \wedge A(x,y,x_1,...,x_n))$$

suit de la directive de définition ontologique.

Ce schéma d'axiomes n'est cependant pas valide en général. En effet, une fois introduit l'axiome de l'infini, on ne pourra développer une théorie générale des ordinaux à la von Neumann sans ajouter l'axiome de remplacement.

L'axiome de l'infini

Cet axiome peut s'énoncer ainsi :

$$(\exists A)(\varnothing \in A \wedge (\forall B)(B \in A \supset B \cup \{B\} \in A)).$$

Avant de le commenter, nous allons réintroduire une notation plus familière en ce qui concerne les ensembles. Lorsque *a* est un individu nous écrirons $\{a\}$ plutôt que $\{\}a$. Si a dénote la multiplicité d'individus b, c, d, nous écrirons $\{b, c, d\}$ pour $\{\}a$. Plus généralement, $\{\}a$ sera représenté par $\{x_1, x_2, ..., x_n\}$où $x_i \, \varepsilon \, x_i$ et $x_i \, \varepsilon \, a$ pour tout x_i.

L'axiome de l'infini dit que l'on peut *ensembler* $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, ...$

La validité de cet énoncé n'est pas une conséquence de l'ontologie élargie. Rien dans les propriétés de $\{\}$ ne permet de dire que cet *ensemblage* est légitime (cela n'est pas surprenant étant donné le second théorème de Gödel !). On quitte ici le terrain de la définition constructive explicite.

Il semble donc que l'ontologie et la théorie des ensembles soient à égalité en ce qui concerne l'existence d'un ensemble infini : elles sont indépendantes de cet axiome.

On remarquera que cet axiome ne vient pas contredire le fait qu'aucun ensemble ne contient un nombre infini d'emboîtements de $\{\}$. L'axiome implique cependant qu'il y a un ensemble qui contient des ensembles formés d'emboîtements finis non bornés.

Quels sont les liens entre les classes distributives et les classes collectives et les ensembles ? Ce sont trois entités différentes. Il suffit de rappeler que si *a* est un individu, $Kl(a), el(a)$ et a sont identiques. Donc ne peut être ni $Kl(a)$ ni $el(a)$.

4. En guise de conclusion

Que peut-on conclure de ce petit exercice ? Premièrement, la théorie des ensembles n'est pas si éloignée que l'on aurait pu le croire de l'ontologie, du moins d'un point de vue formel. Elle l'est cependant du point de vue philosophique. Je me permets de citer Krivine dans son introduction :

> Le point de vue adopté dans ce livre pourra paraître étrange à ceux qui considèrent que la théorie *axiomatique* des ensembles doit être placée au début des mathématiques (ce qui est peut-être vrai pour la théorie naïve). En effet, on ne demande pas au lecteur d'oublier un seul instant ce qu'il sait déjà en mathématiques ; au contraire, on s'appuie essentiellement sur l'habitude, qu'il a acquise, de manier des théories axiomatiques, pour lui en présenter une nouvelle : la théorie des relations binaires qui satisfont les axiomes de Zermelo-Fraenkel.[1]

Adoptant, pour les besoins de la cause, ce point de vue, on peut dire que l'ontologie est elle aussi une théorie naïve qui se présentait comme le début des mathématiques. Sa grande qualité ne ressort pas de sa comparaison avec la théorie ZF : contrairement à cette dernière, l'ontologie, comme la théorie naïve des ensembles, est une théorie qui transporte avec elle une interprétation intuitive. Sa grande qualité est qu'elle ne génère pas la contradiction de Russell car ses ressources de base, les définition thèses et l'axiome de l'ontologie, ne permettent pas la création de cercles vicieux.

Deuxièmement, l'ontologie et ZF sont à égalité en ce qui concerne l'existence d'un ensemble infini et plus généralement la théorie des ordinaux transfinis.

L'ontologie n'a pas eu le succès qu'elle aurait mérité, probablement parce qu'elle est apparue trop tard, apportant une solution philosophique à un problème qui n'en était plus un. La théorie ZF, même si on ne peut prouver sa consistance, offre un cadre suffisamment sécuritaire pour faire des mathématiques la conscience tranquille. On peut cependant se demander, pour terminer sur une spéculation, à quoi aurait pu ressembler le paysage logico-mathématique si l'ontologie était apparue avant la théorie des types et la théorie ZF car d'un point de vue philosophique fondationnel, c'est la plus intéressante des trois.

[1] (Krivine 69, 6)

Références

KRIVINE, J.-L., 1969. *Théorie axiomatique des ensembles*, Paris, PUF.

KÜNG, G. et CANTY, J. T., 1970. « Substitutional Quantifica-tion and Lesniewskian Quantifiers », *Theoria*, 36, 2 : 165-182.

MIEVILLE, D., 2004. *Introduction à l'œuvre de S. Leśniewski, Fascicule II, L'ontologie*. Neuchâtel : Centre de recherches sémiologique.

SIMONS, P., 1982. « On Understanding Leśniewski », *History and Philosophy of Logic* 3 : 165-191.

SOBOCINSKI, B., 1984. « Lesniewski's analysis of Russell's paradox. » In *Lesniewski's systems: ontology and mereology*, edited by Jan T. Srzednicki and Frederick Rickey, 11-44. La Haye : Martinus Nijhoff.